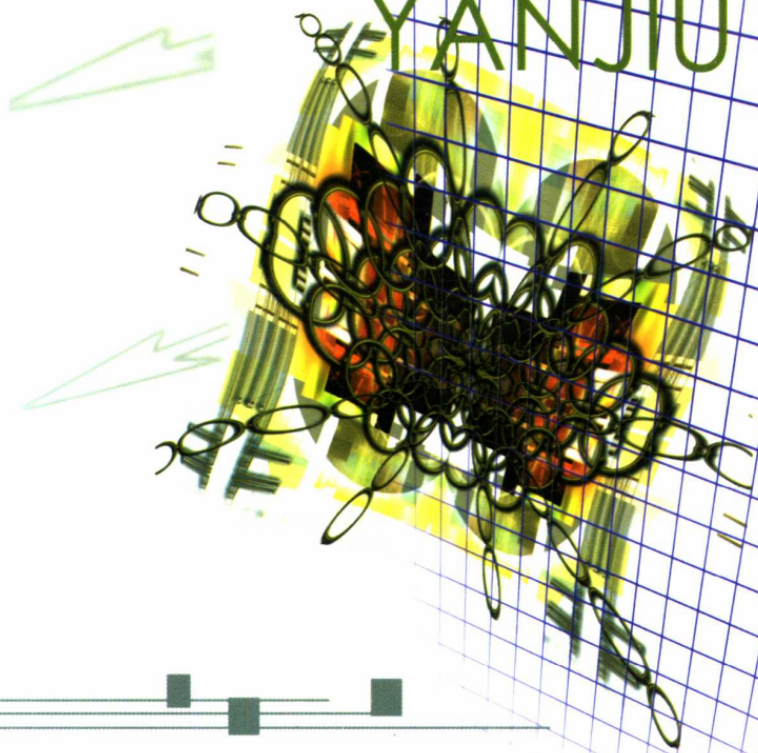


TONGJI  
ZHEXUE  
YANJIU



中国统计出版社  
China Statistics Press



# 统计哲学研究

TONGJI ZHEXUE YANJIU

龙 著

TONGJI ZHIXUE YANJIU

<http://www.stats.gov.cn>

ISBN 7-5037-3939-8



9 787503 739392 >

ISBN 7-5037-3939-8/C. 2005

定价: 16.80 元



# 统计哲学研究

王振龙 著

 中国统计出版社  
China Statistics Press

## (京)新登字 041 号

### 图书在版编目(CIP)数据

统计哲学研究/王振龙编著.

- 北京:中国统计出版社,2002.7

ISBN 7-5037-3939-8

I. 统…

II. 王…

III. 统计学-哲学思想-研究

IV. C8

中国版本图书馆 CIP 数据核字(2002)第 082361 号

### 统计哲学研究

---

作 者/王振龙

责任编辑/刘国宁

封面设计/刘国宁 张建民

出版发行/中国统计出版社

通信地址/北京市西城区月坛南街 75 号 邮政编码/100826

办公地址/北京市丰台区西三环南路甲 6 号

电 话/(010)63459084、63266600-22500(发行部)

印 刷/科伦克三莱印务(北京)有限公司

经 销/新华书店

开 本/850×1168mm 1/32

字 数/170 千字

印 张/7

印 数/1-2000 册

版 别/2002 年 7 月第 1 版

版 次/2002 年 7 月第 1 次印刷

书 号/ISBN 7-5037-3939-8/C·2005

定 价/16.80 元

---

中国统计版图书,版权所有,侵权必究。

中国统计版图书,如有印装错误,本社发行部负责调换。



## 内容摘要

本书立足于哲学基石之上,用哲学的观点对统计的本质、认识系统及其功能、方法背景、逻辑思维等基本问题进行了审视。探索并研究了统计的哲学意义。全文共分五个部分,各部分的主要内容如下。

导言,确立了统计哲学思考的研究方法和研究思路,以及对统计进行哲学思考的现实意义和理论价值。

第一章,统计本体论。从统计最基本、最本质的细胞——统计信息着手,探讨了统计信息的特征和运动过程及其系统结构,在此基础上提出并论述了“统计信息活动是一种认识活动”,“统计的内容是客体的数值规律性”,“统计方法是认识数值规律性的法则”,“统计学是一门亚哲学”等观点。

第二章,统计认识论。首先运用认识论的观点研究了统计认识的基本属性,统计认识运动的特殊性,基本的统计认识观,详细地阐述了统计是一个不断循环的再认识过程的观点,深入地分析了统计认识的客观性和相对性的矛盾运动。其次,以系统论为指导,就统计认识系统的各个组成要素即统计认识主体、统计认识客体、统计认识手段和统计认识宿体,分别剖析了其质的规定性、系统结构及其在统计认识系统中的基本功能。

第三章,统计方法论。这里对统计方法的研究,既不是论证统计方法的数学原理,也不是系统地建立统计的方法体系,而是聚焦于统计方法的哲学原理。一方面,通过对统计方法的普适性的研究,进一步证明统计科学客观上具有“亚哲学”的地位,另一方面,对一些统计方法展开了哲学层面的分析,探索和揭示

了这些方法的哲学背景。

第四章,统计思维。提出了统计思维的概念并论述了其特点。归纳和概括了统计推理的范式,用批判的观点辩证地揭示了统计推理的不彻底性,并对一些具体的统计推理过程进行了较为深入的探究。

## **Abstract**

This paper based on the foundation of philosophy, examines such basic problem as system and its function, the background of method and logical thinking of statistics from the point of view of philosophy, and probes into the philosophical sense of statistics. It contains five chapters, and the main contents of each chapter are:

### **Introduction**

This paper establishes the research method and train of thoughts of philosophical thinking of statistics, and theoretical value of philosophical thinking if statistics.

### **Chapter one**

Statistical Ontology. Starting with the most fundamental and essential cell of statistics——statistical information, it explores the characteristic and motion process of statistical information and its systematical structure. On this base, it puts forward and discusses the views as to: “statistical information activity is a kind of cognitive activity”, “the content of statistical is the numerical regularity of objects”, “statistical method is the law of cognitive regularity”, “statistical is a subject of sub-philosophy”, etc.

### **Chapter two**

Statistical Epistemology. First, it studies the basic attribute

of statistical cognition, the specialty of statistical cognition motion and the basic statistical cognition concept from the point of view of epistemology, and elaborates in detail the point that statistics is a endless circulatory recognitive process and deeply analyses the contradictory movement between objective and relative statistical cognition. Second, under the guidance of systematic theory, it dissects respectively the formula of essence, systematic structure and their basic functions in the statistical cognitive system as to each component factor of statistical cognitive system, that is, statistical cognitive subject, its object, method and its cognitive root.

### **Chapter three**

Statistical Methodology. The research of statistical method here is neither the mathematical principle of the verification of statistical method, nor establishing systematically the method system of statistics, but focus on the philosophical principle of statistical method. On one hand, it further certifies the position of "sub-philosophy" of statistics objectively through the study of extensivity of statistical method. On other hand, it analyzes some of the statistical method in the view of philosophy, and explores and discloses their philosophical background.

### **Chapter four**

Statistical thinking. It puts forward the concept of statistical thinking and discusses its characteristics. It summaries and outlines the normal form of statistical inference. From the point view of criticism it dialectically discloses the incompleteness of statistical inference, and also does some further research on a few specific statistical inference processes.

# 目 录

导言 .....	( 1 )
第一章 统计本体论 .....	( 7 )
1.1 统计实践:一种信息操作 .....	( 7 )
1.1.1 统计信息 .....	( 7 )
1.1.2 统计信息运动过程 .....	(17)
1.1.3 统计信息系统 .....	(20)
1.2 统计信息系统:一种认识运动 .....	(23)
1.2.1 信息是统计认识的中介体 .....	(23)
1.2.2 统计认识是客体信息的一系列变换 .....	(24)
1.2.3 统计信息活动是一种认识活动 .....	(26)
1.3 统计内容:客体的数值规律性 .....	(27)
1.3.1 统计从量的方面把握客体 .....	(27)
1.3.2 统计揭示客体的统计规律性 .....	(28)
1.4 统计方法:认识数值规律性的法则 .....	(30)
1.4.1 统计方法的总原则是大量观测基础上的 求同存异 .....	(30)
1.4.2 统计思维是一种系统思维 .....	(32)
1.5 统计学:一门亚哲学 .....	(33)
1.5.1 统计学的发展史是人类认识世界的历史 .....	(33)

1.5.2 大量观测和实验是统计学产生和发展的基石 .....	(34)
1.5.3 统计学与实践 .....	(35)
1.5.4 统计观点作为哲学意识形态的地位 .....	(39)
<b>第二章 统计认识论</b> .....	<b>(48)</b>
2.1 统计认识一般 .....	(48)
2.1.1 统计认识运动 .....	(49)
2.1.2 统计认识的基本属性 .....	(50)
2.1.3 统计认识观 .....	(55)
2.1.4 统计认识过程 .....	(57)
2.1.5 统计认识的客观性和相对性 .....	(66)
2.1.6 统计认识系统 .....	(68)
2.2 统计认识主体 .....	(73)
2.2.1 统计认识主体的质的规定性 .....	(73)
2.2.2 统计认识主体的形式 .....	(75)
2.2.3 统计认识主体的结构 .....	(77)
2.2.4 统计认识主体的功能 .....	(80)
2.3 统计认识客体 .....	(82)
2.3.1 统计认识客体的属性 .....	(82)
2.3.2 统计认识客体的形式 .....	(88)
2.3.3 统计认识客体的结构 .....	(89)
2.3.4 统计认识客体的功能 .....	(93)
2.4 统计认识手段 .....	(94)
2.4.1 统计认识手段的基本属性 .....	(94)
2.4.2 客体信息场 .....	(95)
2.4.3 主体生理结构 .....	(97)
2.4.4 主体认知结构 .....	(99)
2.4.5 统计认识工具 .....	(100)
2.4.6 统计认识方法 .....	(102)



2.4.7 统计认识的信息符号 .....	(105)
2.5 统计认识宿体 .....	(108)
2.5.1 统计认识宿体的界定 .....	(108)
2.5.2 统计认识宿体的功能 .....	(109)
<b>第三章 统计方法论</b> .....	(111)
3.1 统计方法一般 .....	(111)
3.1.1 统计方法的普适性 .....	(111)
3.1.2 统计认识方法的主观方面和客观方面 .....	(116)
3.1.3 统计认识方法的突出特点 .....	(119)
3.2 统计信息建构方法 .....	(125)
3.2.1 客体的观念建构 .....	(125)
3.2.2 统计观测方法 .....	(131)
3.2.3 统计实验方法 .....	(135)
3.2.4 统计测量方法 .....	(138)
3.3 统计信息加工方法 .....	(144)
3.3.1 统计描述方法 .....	(144)
3.3.2 统计推断方法 .....	(158)
3.4 统计信息利用方法 .....	(166)
3.4.1 统计解释 .....	(166)
3.4.2 统计预测 .....	(170)
3.4.3 统计决策咨询 .....	(173)
<b>第四章 统计思维</b> .....	(174)
4.1 统计思维的特点 .....	(174)
4.1.1 统计思维一般 .....	(174)
4.1.2 经验思维与理论思维的统一 .....	(176)
4.1.3 直觉思维与抽象思维的统一 .....	(177)
4.1.4 形象思维与逻辑思维的统一 .....	(178)
4.1.5 归纳与演绎的统一 .....	(179)
4.1.6 探索与实证的统一 .....	(180)

4.1.7 实践与理论的统一 .....	(181)
4.1.8 历史的与逻辑的统一 .....	(182)
4.2 统计推理范式 .....	(183)
4.2.1 基本范式 .....	(183)
4.2.2 统计推理的不彻底性 .....	(186)
4.3 统计推理 .....	(188)
4.3.1 统计推理一般 .....	(188)
4.3.2 统计演绎 .....	(190)
4.3.3 统计归纳 .....	(191)
4.3.4 统计类比 .....	(194)
4.3.5 贝叶斯推理 .....	(197)
4.4 统计解释及其逻辑模型 .....	(206)
4.4.1 统计解释 .....	(206)
4.4.2 统计解释的逻辑模型 .....	(206)
<b>参考文献</b> .....	(209)
<b>后记</b> .....	(211)

# 导言

## 一、什么是统计哲学思考

在统计的发展长河中,统计人关于统计理论、方法及其在各个领域的应用,都有了较深入的研究,形成了较为系统、完整的统计学分支学科,而对统计的哲学基础研究却甚为罕见。统计哲学思考就是对统计的思想进行哲学的考察。统计的思想蕴含在统计方法之中,统计方法是统计思想的载体,统计思想是统计方法的实质。也可以说,统计方法是固化了的统计思想,是实现统计思想的程序。在这个意义上说,统计思想就是现实的一切统计方法。另一方面,哲学是人类理性关于宇宙的总看法的一系列观念、概念体系。我们对统计思想进行哲学思考不是给现实的统计方法贴上哲学概念的标签,也不是强行使现实的统计方法服从于某一哲学思想,或者对统计方法作出哲学解释,而是探索统计思想的哲学意义,归纳、概括和抽象统

计方法的思想并使之升华到哲学层面。从而,在哲学层面上研究统计的本质、对象、方法体系以及学科体系等理论问题。同时,在思考的基础上进一步深入研究并使之理论化、系统化,将会产生一门新兴学科——“统计哲学”,促使统计学的发展发生质的飞跃,并充实和丰富哲学的内容。

## 二、为什么要对统计进行哲学思考

统计实践呼唤着哲学思考。统计从其萌芽时期的“结绳记事”开始,就执行着计数的功能。随着人类的行为不断从“自发”向“自觉”转变,人类认识世界的需求越来越广泛、越来越迫切,认识的程度也越来越精细。在这一强大动力的作用下,统计的计数功能逐渐向科学地获得足够的信息并据此探求客体数量结构、运动变化规律的认识功能和方案优化、客体运行的信息监测、反馈、精选调控方案的改造功能发展。其发展轨迹从纵横两个方向展开,使之不断得以深化。横向认识的深化过程可以概括为:特征描述(个别指标)→总体状态描述(指标及其体系)→总体分布特征(位置、形状)→结构分析(各种要素及系统与环境间的关系)→要素分析→环境分析→未来走势及状态推断→调控方案。纵向过程总的来说是沿着:历史轨迹→现在跟踪→未来发展的时间序展开的。从认识的深度来看往往表现为:发展水平(状态)→一般水平(平均水平)→发展变化(方向及速度)→时间影响(与时间的关系)→分量分析→平稳性分析→非平稳性分析→奇异点及转折点分析。从统计的应用领域来看,已经覆盖了自然、社会和思维领域,发展成为认识世界、控制实践活动和进行科学研究的重要工具之一。不仅如此,统计的应用已渗透到人们日常生活的方方面面。不具备一定的统计常识,无法正确解读天气预报的降水概率;不具备一定的统计常识,就理解不了与人们生活息息相关的物价指数;不具备一定的统计常识,

就不能正确掌握和利用形形色色的“市调”和“民调”;……

特别是在现时代,信息和通讯技术相结合产生了信息网络技术,并以惊人的浸透力和生长力迅速覆盖人类活动的各个领域和各个环节。人们对信息采集、过滤、加工和利用的需求更为迫切,甚至成为必须;计算机强大的知识存储功能和计算功能使得人们可以抽象掉统计方法的深奥的数学理论和复杂的运算,从而进一步拓展了统计的应用空间,并使统计认识方法的应用也将发生新的变化,即定量分析将变得相对容易,而定性分析将成为统计认识的主要方面。

与此同时,我们必须实事求是地承认,在现实中误用统计者大有人在。其中既有使用失误者,也有“别有用心”者,不能正确解读统计者则更比比皆是。以致于人们戏称统计为“数字游戏”、“造假机器”、“诡辩工具”、“科学谎言”等等。因而,统计的本质是什么?统计认识系统的结构和功能是什么?统计的方法论基础是什么?统计的逻辑推理过程怎样?以及怎样才能正确使用和解读统计等等,这一系列问题亟待探究并作出回答,方能还统计之本来面目。

统计学的发展要求进行哲学思考。自从威廉·配第的《政治算术》一书出版,宣告统计学诞生以来,随着人类对定量认识需求的不断增强和统计实践的不断发展,统计学的研究和发展从两个方向进一步深入展开。一是统计方法的研究,二是结合具体认识对象特性的统计应用研究。

在统计方法研究方面,围绕统计指标设计、统计信息采集、数据显示、分布特征和规律、结构刻画、依存关系测定、因素分析、动态发展规律的探索、样本设计、统计推断、统计预测、统计决策以及模糊对象的统计方法等课题展开,形成了门类众多的统计方法分支学科。当今,信息网的产生和发展对传统的统计认识方法产生了巨大的影响。信网调查,比赋模拟等方法将成为具有现时代特征的重要的统计认识手段。

在统计应用研究方面,统计方法与关于具体应用对象的实质性科学相结合,产生了各种各样的应用统计学,如物理统计学、天文统计学、卫生统计学、人口统计学、经济统计学、环境统计学、心理统计学,以及目前正在研究的可持续发展统计学等等。

此外,还有研究统计实践和统计科学发展的统计史学以及规范统计行为的统计法学。

这些形形色色的统计学各自具有独特的研究对象和特有的研究方法,同时又相互交叉、渗透,因而引起了“流派”之争、“门户”之战,严重地影响了统计科学的发展。这些争论的一个主要理论缘由就是争论者忘记了“白马非马”这一哲学原理。“白马”认为自己是“马”,因而不承认“红马”、“黑马”是“马”;而“红马”则强调自己是“马”,而不承认“白马”和“黑马”。近几年来进行的“大统计学”大讨论和研究,实质上是关于“马”的讨论,是在各种“有色马”中寻找“马”。

更为重要的是,长期以来各种统计学的研究都停留在“怎样操作”这个技术层面,尚未上升到“为什么要这样操作”以及“这样操作与客观的相合程度”这个理性领域。“方法学”侧重于数学推理;“应用学”则致力于与对象结合;“史学”集中于“过程”;“法学”强调统计行为。统计科学的发展必须向哲学理性升华,经过对统计学各分支学科的本质、共性、规律进行理论概括、抽象,形成统计本体论、认识论、方法论和统计逻辑,然后在这些统计的哲学观念、哲学思想的指导下,进一步向统计理论的广度、深度发展和完善。只有这样,才能突破争论的疆局,只有这样,才能建立起真正意义上的“大统计学”。也就是说,“大统计学”是“马”,而不是各种“有色马”的集合。

### 三、怎样对统计进行哲学思考

统计哲学思考是一个巨大而深沉的课题。它既包括对统计



实践的概括,又包括对统计理论的反思;既包括对统计发展过程的历史考究,又包括对统计现实的探析;既包括对渗透到各种统计分支学科中的统计思想进行提炼和升华,也包括对其疏漏和欠缺进行批判;对于研究者而言,对统计进行哲学思考,既要具有足够的哲学知识准备,又要准确地理解和把握各门统计分支学科的思想。

然而,对统计的哲学思考又应是自由的。这首先表现在对统计进行哲学思考必须始终坚持“百花齐放,百家争鸣”的方针。必须允许存在不同的分歧,必须允许有不同的见解,既不能强制,也不能压服。这是科学地进行哲学思考的前提,也是科学发展的内在要求。其次,对统计进行哲学思考是一种学术探索,是一种创新,不唯书本,不唯权威。再次,对统计所进行的哲学思考,允许存在尚不完善、甚至是错误的观点,因为哲学就是批判的哲学,就是在批判中不断得到发展的。最后,对统计进行哲学思考既可以全面出击,面面俱到,也可以将精力集中到某一个、或某几个问题深入思考。正是以这些原则为基石,几年来,作者大胆地在统计与哲学的结合部进行了一些探索性的思考和研究。

在对统计进行哲学思考时,首先碰到的问题是以何为切入点,应当从何处着手去再现统计的历史。根据辩证法的要求,作者首先研究了统计之本体,紧接着对其最基本的认识功能展开了探讨,并在此基础上进一步研究了统计方法论和统计思维问题。

为了揭示统计的本质,必须再现其发展的现实历史过程。但是,这里值得注意的是只有在对象的本质已为我们所知的场合下才能再现对象发展的现实历史过程,这就是辩证法。比如,要研究人的本质就得先知道人类产生和发展的历史,但是要研究人类的历史又需要具有关于人的本质的一定知识,否则就可能把先于人类的动物(猿)也看成人了。同样,这里对统计发展

史的研究,主要着力于保持现实历史过程的基本线索上,而不是具体环节上。鉴于这样的认识,这里在研究统计的发展历史时,是从现时、从统计最成熟的形式和发展阶段开始研究,这时,统计的本质得到充分的体现,不为与它无直接关系的偶然性所掩盖。在研究统计发展最高、最成熟阶段的基础上提出统计的最初规定。这些最初规定带有抽象性,不够深刻,但作为研究统计历史发展过程的线索是必要的;它们是研究统计历史过程的出发点,因为它们在某种程度上反映了被考察客体的形式与发展过程。这一点无须置疑。

# 1 统计本体论

## 1.1 统计实践：一种信息操作

### 1.1.1 统计信息

根据辩证法的要求,研究对象的本质必须从对象发达形态中最简单、最大量、最常见的东西开始,并且要从其在萌芽状态中就包含有复杂的发达形态的全部丰富性和突出特点的这样一种简单的东西开始。因此,我们研究统计之本体也必须从其最基本、最本质的细胞——统计信息这一概念开始。统计信息是信息的一种。为此,我们从信息一般来探究统计信息的具体概念、特征及其运动。

#### 一、信息世界

关于信息的概念从不同角度考察有多种含义。信息作为日常用语,是指信号、消息。作为科学术语则是指数据、消息中包含的、消息接收端预先不知道的定义。在不同学科中,信息的

涵义也有所不同。在信息论中,信息使消息中所描述事物的不确定性减少,如不提供信息,不确定性就大。可见,信息是用来减少人们无知程度和不确定因素的媒介。在通信和控制系统中,信息则是对世界上一切事物和现象的共性——相互联系的复杂性、有序性和差异性的高度概括,是物质普遍联系的一种形式。作为哲学范畴的信息概念也尚未形成统一的定论,本文认为信息是标志物质间接存在性的哲学范畴,是同世界的物质、能量紧密联系在一起普遍属性,是物质存在方式、运动状态以及相互关系的自身显示。信息不是事物本身但它反映了事物的性质、状态以及规律性。信息已成为表示客体的变化和客体之间相互差异或关系的重要标志。有着重要的认识论和方法论意义。从认识论来看,信息是认识的手段,又是认识的结果,而认识则是信息过程的一种形式。从方法论来看,信息是反映客观物质世界的一种重要思维形式。信息可分为结构信息和相对信息。前者给出了系统过程本身内部情况的描述,是系统的有序程度的度量即负熵,而后者则是与两个过程的关系联系在一起。如果一个对象受到了另一个对象的作用,并且发生了反映,那么,这种作用的变化在控制系统中第一个对象便成为第二个对象的信息携带者。第一个对象的这种变化正是作为另一个对象的信息携带者使控制系统得以发挥功能。从这点出发,由于物质相互作用的普遍性和物质、时间的无开端性,任何物体都已经将自身溶化成了具有特定结构和状态的凝结着种种信息的信息体了。正是物质的这种信息体性规定着任何物体都是一个直接存在和间接存在的统一体。客观决定了在物质世界之外还存在着一个信息世界。它是物质和精神的中介。

可见,信息是人类认识发展过程的基本支点和人类探索客观事物规律进程中的一个阶段,成了人们认识和改造世界的方法和工具。从本质上来看,信息反映了物质之间相互联系、组成程序以及它们各自的差异性,是物质的一种属性或高级运动形

态。它既不同于物质的实体,也不同于物质的运动,但又离不开物质和运动。

作为物质的间接存在,信息既包括客观存在而不实在,即客观间接存在,也包括主观存在即主观间接存在。物质的间接存在性,归纳起来无非是三个方面:一是关于自身历史的反映(包括曾经发生过的与它物的关系);二是关于自身性质的种种规定,这些规定在其展示的时刻是一种直接存在的过程,但只在其未展示的时候还只能是一种现实的间接存在;三是关于自身变化、发展的种种可能性,这便是关于事物历史、现状、未来的三种间接存在。这三种间接存在具体凝结在一个具有特定结构和状态的直接存在物中。任何物的直接存在的结构和状态都是由它所凝结着的间接存在所规定的;任何物的结构和状态都映射和规定着关于自身的历史、现状和未来的信息。

信息可以游离其客体而被记录、存储、传递,加工和利用。但是,不论是发送信息的信源,或是传递信息的载体、信道,还是接受信息的信宿都是物质。信息必须通过物质才能存储、传播。

## 二、情报、知识、信息

知识是一种特定的人类的信息,是整个信息的一部分。在一定的历史条件下,人类通过有区别、有选择的信息,对自然界、人类社会及思维方式和运动规律的认识与掌握,并通过大脑的思维使信息系统化,形成知识。社会实践是知识的源泉,信息是知识的原料,知识是系统化的信息。可见,信息是现象与知识的中介,是事物运动状态或存在方式的直接或间接的表述。信息既不等于现象,也不等于知识。信息与现象的区别在于信息具有知识的秉性,而信息与知识的区别在于信息是知识的毛坯,没有经过加工、提炼、升华。

情报是信息的一部分,也是知识的一部分。情报是针对某种社会目的需要及时传递的新知识的信息。知识包括新旧知识两部分。一切新知识首先存储于人脑中,成为主观知识;然后,

再通过某种物质载体,将主观知识记录下来,变成可以传递交流的客观知识;再围绕某种社会目的交流、传递知识,便成为情报活动。当一种新知识被一种更新的知识取代之后,便成为普及教育的常识,就不再是情报的范围。信息包含知识,知识包含情报,即

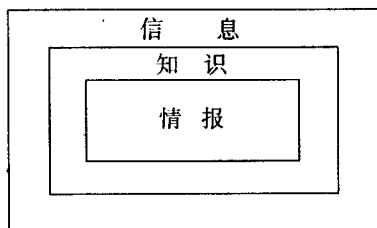


图 1 情报、知识、信息关系示意图

### 三、统计信息的概念

在这里需要特别强调的是统计认识系统中的统计信息概念不同于信息论的统计信息。在信息论中,统计信息是指抽象掉信息的语义和实效,而只考虑客观的不定度。严格地说是指申农关于信息的统计模型。在这里统计信息被定义为:利用统计方法获得的信息。具体地说,统计信息是统计主体根据宿体的需要,运用统计方法,经过一系列统计活动所取得的关于被认识客体的状态、结构、功能及其发展变化规律性的客观描述。它是一种依据方法定义的信息,即是对信息总体按其获取方法不同而加以标识的结果。其内涵在于获取信息的手段是统计方法,其外延不受时空范围以及统计认识主体(即信息来源的渠道)的限制。不论是客体的历史运动轨迹、规律或现在状态的描述,还是未来行为的揭示;不论是来自自然现象、社会现象和思维活动,还是科学研究或管理过程;不论是由专门从事统计活动的政府统计机构、企事业单位的统计部门或民间统计组织,甚至是个别统计工作者,还是由科研院所以及各个活动部门;不论是宏观或中观,还是微观;只要是按照统计规则、运用统计方法获得的



信息,都是统计信息。

统计信息既包括所研究的整体信息,也包括其子系统及元素的信息;既包括状态、规律、水平的原始信息,也包括结构、特征、行为、功能、规律性的深层次信息;既包括历史、现状的描述信息,也包括未来的预测信息;既包括认识世界的分析信息,也包括改造世界的决策支持信息。

显然,统计信息与统计数据是两个性质不同的概念。统计数据是统计信息的一种载体,是载荷统计信息的一种符号。可见,二者是内容与形式、目的与手段的关系。统计数据包含了有系统的全部信息,但要成为有用的知识,还必须经过开发。另一方面,失去了统计数据,那么相应的统计信息也就随之而消失。这就是统计工作之所以要采集、存储、加工处理统计数据的原因。统计数据是统计信息的一种基本物理载体。除了统计数据之外,统计信息还有其他载荷形式,如统计模型、文字记录、统计图表等。

我们说统计是信息测量和处理的手段,是在认识论上而言的。与通讯理论中的信息度量的含义有很大的区别。在通讯理论中,信息实质上仅仅指语法信息层面的概率信息,所谓的信息度量就是对随机信息的概率描述,是信息的不定度的测定。它抽象掉了信息的语义和语用,抽象掉了信息传输中的物理形式,抽象掉了接收者的主观方面。统计信息理论中的信息,是作为认识的中介而言的,是作为统计认识主体运用统计方法对客体的属性、关系、运动状态、运行规律等方面进行的量测和记录,并尽力使其与客体保持同构,在此基础上经过一系列的同构变换,突现出便于宿体把握的客体的结构和特征。可见,在统计信息系统中,始点是客体即信源,而终点是用户即信宿,统计主体是中介。这样形成了一个完整的认识系统。统计主体的作用在于不仅要获得与信源(客体)结构和有序性一一对应的原始信息,而且还要根据自己的认知结构,并结合环境约束,经过一系列加

工处理和主客体信息交流突现出客体的结构特征,最后提供给信宿用于控制和改造客体。

#### 四、统计信息的特性

1. 数据性。统计信息的数据性表现在两个方面。一是就其载体而言,统计信息主要是用数据载荷信源(即统计客体)的状态、结构、动态特征及发展变化规律性。数值化是统计信息本身所具有的特殊性。即统计依据一整套科学的统计指标体系,通过对大量的经济、社会和科学技术或自然现象进行系统观测、估计等统计手段,提取有序的、有规律的数量关系和数量界限,得到反映所研究系统及其子系统的变化特性及规律性的数量性信息。因此,统计信息的表现形式主要是统计数据(包括描述、判断、推断、预测等数据)。当然,也有文字报告等其他载体。二是就统计信息的内容而言,统计信息反映的是客体的数量方面,主要包括数量多少,现象之间的数量界限。统计信息内容的数量性决定了统计信息主要以统计数据为载体。这是统计信息的基本特征。

2. 社会性。统计信息的社会性是由统计信息的来源和使用二者共同决定的。统计信息是随着人类的生产而生产的。在当今社会,凡是有人类活动的地方都有统计,大到一个国家,小到一个企业,统计无所不在;另一方面,就某一系统而言,统计涉及各个方面和各个环节。如工业企业生产经营活动统计,既包括供、产、销各个环节,也包括人、财、物各个方面;既包括成果,也包括耗费;既包括系统本身,也包括该系统与他系统之间的数值联系。也就是说,统计信息来自人类一切活动的各个领域、各个方面、各个环节,是关于各个人类活动的知识。这是统计信息区别与其他信息的一个重要方面。从广义的角度来看,凡是用统计数据反映的社会各方面的情况都是统计信息。因此,统计信息的信源是人类一切活动,统计信息的内容是全社会的。从这个意义上来说,就整体而言,统计信息是认识人类活动的一个完

整的知识体系。

由于统计信息源于全社会,所以统计信息是全社会的财富,属全社会所有。统计信息的宿体(即信息使用者)是一切人类活动的主体。只要在国家保密规定允许的范围内,一切个人、机关、团体都有权使用统计信息。统计信息应服务于人类认识世界和改造世界的一切活动。

3. 时效性。统计信息的时效性有两种涵义。一是及时性。所谓及时,首先是随着现象的发展变化,及时搜集、提取统计信息。统计信息是客体存在方式和运动状态的反映。统计客体——物理系统是在不断地运动、变化和发展的,随着时间的推移,随时都有大量的新信息不断产生。因此,统计主体要不断掌握新情况、新问题,适应客体的发展变化,及时收集和传递新的统计信息,更替和补充已经积累的统计信息,保持统计信息的时效性。其次,从信息的实质来说,它是“使用者预先不知道的定<sub>义</sub>”。人们无不想变不知为知,在这个意义上,一则迟到的,已为人知的情况、消息、数据就不能称为信息。常言道:“有钱难买早知道”,这充分反映了人们希望尽早地获得其所需信息的迫切心情,同时,从另一侧面说明了信息的价值在于及时。一则及时的信息可使人们尽早地认识系统的状态、工作的业绩、未来的趋势、存在的问题、应采取的措施,从而适时调控系统的行为,使之永远保持最佳的结构,协调运行,很好地实现系统的目的。现实中,一则及时的信息使多少经营者成功,而一则迟到的信息又使多少决策者失败,实际例子,枚不胜数。

这是统计信息作为信息的一种,所具有的一般信息的时效性。统计信息作为一种特殊的信息,在“时效性”上又有其自身的特殊意义。统计信息是人类活动过程的客观反映,是人们认识不断深化的依据。既然是个过程,统计信息的使用价值除了满足人们认识系统在某一特定时刻上的状态,从而作出正确的决策,有效地调控系统的行为而外,还必须满足人们认识系统动

态结构和发展规律的需要。这就决定了统计信息除了具有现实的价值之外,还有历史价值。这就是统计信息时效性的第二种含义。统计信息的历史价值决定了储存统计信息的必要性。系统地存储起来的关于客体历史行为的统计信息是人们研究和认识客体发展变化趋势和规律的重要依据和必要前提。

4. 失真性。失真是指人们在提取、计量、处理、传递以及开发利用信息等过程中,由于噪声的干扰,使得到的统计信息与信源所发出的信息不一致的现象。失真是任何信息都具有的共性,对于统计信息来说,更具有特别重要的意义。

统计信息失真的类型主要有倾向性的信息失真和非倾向性的信息失真。所谓倾向性的信息失真,是指人为的、有意造成的失真,如故意谎报、有意歪曲和篡改,造成大面积的、长期的、严重的失实。非倾向性的信息失真是指在统计设计和统计调查过程中由于技术上的失误而引起的统计数据的失真。如指标设计、分组设计的误差、计量误差等等。同时,在统计整理过程中,进行手工汇总过程的记录、计算、抄写、打印等产生的误差以及机器汇总过程中,编码、录入、数据处理、打印过程产生的噪声,都会使统计数字失真。

5. 客观性。统计信息的客观性有绝对与相对之分。所谓绝对意义是指统计信息是统计客体数量方面的客观反映,是所研究系统的状态、结构及运行行为的客观记录,是系统运动规律的客观描述。统计信息是统计客体的客观存在,而不是客观实在,也不是主观存在。统计信息客观性的相对性是指统计信息相对于失真性以及其它信息来说,具有客观性。在利益主体多元化的条件下,由于利益的驱动,在搜集和量测信源所产生的信息时,不可避免地会受到来自各个方面的噪声的干扰,使得信息存在一定程度的失真,但由于统计认识主体既不是决策单位,也不是执行部门,而是专门从事信息情报的专业部门。它的职责是依据法律的规定、独立进行统计调查、整理、加工分析并提供、发

布统计信息。它置身于决策的执行系统之外,没有直接的利害关系,可以说比较超脱。所以,统计主体只对信息负责、并尊重实事求是的原则。从而使得统计信息与其他信息相比较,受到的干扰较少。此外,统计信息系统还有一系列科学的信噪监测和净化手段,有统计法制保障,因此统计信息虽然存在着失真,但相对来说,较为客观。

6. 多次使用性。统计信息作为统计部门的产品,具有自己区别于其它产品的特点。其它的一般物质产品,在其消费过程中改变了原来的形态和存在形式,只能满足人们一次消费。而统计信息产品,属于概念的范畴,无论是以有偿服务还是无偿服务形式提供消费者使用,都不会消灭原来存在的信息,即一则信息可以供不同用户使用,满足其各自不同的需要。同一则信息,既可供本部门用于管理活动过程,也可供上级各层次用于宏观调控;既可作为最终产品直接用于具体活动,也可作为进一步开发深层次信息的资源;既可供管理者和决策者管理活动,也可供监督部门监督活动;既可用于描述某时刻系统所处的状态,也可作为研究规律、预测未来的依据;既可用于调控系统现在的行为,也可用于研究系统发展变化规律的历史依据,等等,具有多次使用的特征。

7. 共享性。统计信息作为一种社会资源是广大统计工作者和各部门共同劳动的结果,应全社会共享。统计信息不因一次使用而消失的特性为统计信息的全社会共享提供了可能。一般来说,借助于有形的信息载体,就可以将信息储存起来。因而,从时间上说,统计信息可以多次使用开发;从空间上来说,一则统计信息并不是某个使用者所能独自占用或使用后即消失的,它可供多个用户同时使用,是可以为公众服务的一种资源。统计信息共享不仅包括作为统计的最终产品,供全社会用于认识世界和改造世界,而且包括作为统计的中间产品,供所有需要者用于开发深层次的信息。

8. 再生性。统计信息是有层次性的,最基本的信息就是描述客体在一定时空条件下所处的状态的信息。这些信息,是我们最初从信息源提取和测量的结果。所以称之为原始统计信息。每一则原始信息一般只反映客体某一方面的基本特性,因而也称为描述性统计信息。这是一切统计信息的源泉。在此基础上,运用科学的统计方法对那些具有内在联系的原始信息进行科学的加工分析、类比、综合、归纳等,便可得到不同的反映客体综合特征的新信息。进一步科学处理和研究,便可得到反映客体之间存在的数量联系,质量互变的数量界限以及发展变化规律性的深层次的信息,据此可推知未来时刻的状态,便得到预测性信息乃至决策支持性信息。一般来说,原始统计信息是直接测量的结果,而较深层次的统计信息都是原始信息的再生信息。各类信息的生成过程如图所示。

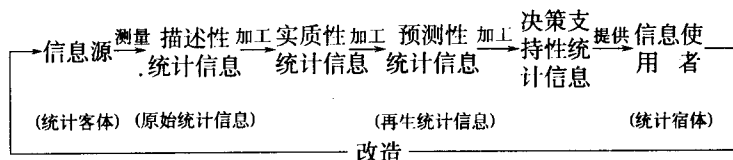


图2 统计信息的生成过程

不同原始信息的不同的科学组合,便会产生不同的再生信息。统计信息之所以会再生,首先是由于统计原始信息是客体行为的客观描述,客体的状态、结构性质及变化规律等属性全都寓于系统化的原始统计信息之中。就象我们通过对大量个性的研究,可以获得共性一样,我们通过对原始信息的研究,就可以获得新的信息。其次,统计信息涵盖之广,涉及各个方面,而现象之间又存在着一定的固有联系,自然这种联系具体地体现在原始信息之中。因而,统计信息之间存在着一定的客观联系,反映信息之间联系的信息,便是再生的新信息。

9. 价值性。统计信息作为统计主体的成果——“统计产



品”,在社会主义市场经济社会中,也是一种“商品”,也具有价值和使用价值两重属性。首先,就统计信息的价值的形成来说,一方面,每则统计信息从收集、整理、加工、计算到传递、反馈等各个环节都要耗费大量的体力劳动和脑力劳动。因而每一则统计信息都凝结着人类大量的活劳动;另一方面,统计活动的进行,必然要占用大量资金、物质装备和科学的信息处理工具,因而每一则统计信息都会有一定的物化劳动。统计信息产品中凝结的人类一般劳动(活劳动和物化劳动)便形成了统计信息产品的价值。其次,就统计信息的使用价值来说,统计信息是一种“特殊商品”,其使用价值就在于满足人们认识世界的需要,使人们变不知为知,由知之甚少到深入了解,由现象到本质,由状态到规律。一旦正确使用,便可产生巨大的生产力,获得经济和社会效益。

统计认识和统计实践是同一过程的两个方面,统计实践的主体也是统计认识的主体、统计实践的客体也即是统计认识的客体,统计认识包含于统计实践之中并从统计实践产生。统计实践的主体和客体之间有着感性现实性的关系,而统计认识的主体和客体有着反映和被反映的关系。在人类认识史上,原来是哲学家研究的问题,后来逐渐分化出来为统计学家研究。经过这个演变过程,统计认识才概括出人类认识客观规律的理论体系,从而独立于哲学之外。

### 1.1.2 统计信息运动过程

就通讯信息系统而言,信息的运动过程是指一则信息由信源到信宿的传递过程。一般来说,顺序经过以下几个环节。

编码。就是信息发出者将其所要发出的信息转换成一定的信号。信息是概念的范畴,要想传输,必须借助于一定的物理载体。所以编码是信息运动的首要环节。编码的基本原则就是客观、简单、易于传输和理解。

传递。即通过一定形式的信道,将信号由信源传到信宿。

译码。就是信息接收者将收到的信号译成信息。

可见,通讯中的信息系统是由信息源、编码器、信道、接收器和信宿五个要素构成的。通过各要素之间的协同,将信息从信源传递到信宿。各要素的基本功能及信息的运动过程可图示如下:

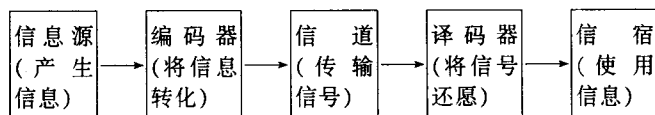


图3 通讯系统中信息的运动过程

统计信息作为统计活动的产品,有其独特的运动形式。它既是一种信息运动,也是一种产品运动。就产品运动而言,任何产品的运动都是一个从生产到消费,由形成到消失的过程。比如,工业产品的运动过程,一般经过以下环节:

第一,产品设计。即根据社会需要和要求,对产品本身的质量、技术、性能、生产工艺、生产流程以及产量、批量、出厂时间、生产成本、目标市场等进行统筹安排。

第二,采购原材料。有计划、有组织地采购价廉物美的原材料,以备生产使用。

第三,原材料的整理和储存。采购回来的原材料不可能全部一次投入到生产过程,特别是同种原材料往往供多种产品使用,因而必须对原材料进行严格的质量检查,并对次品和不合格品进行加工修理或清除处理,然后按品种、性能、用途、规格、质量等有条理地存储起来,随时向生产过程提供所需的材料。

第四,产品生产。根据产品设计,按照生产工艺要求,运用一定的生产工具对原材料进行加工,生产出有一定使用价值的产品。

第五,产品销售。产品生产出来以后,需经过销售使产品从生产领域进入消费领域。

第六,消费。即用户为满足生产或生活需要而使用产品。一方面,消耗掉该产品,用户的需要得到一定程度的满足,结束

了该产品的运动过程；另一方面，又诱发出新的需要，从而形成了新的生产动因，于是开始了又一个新的生产过程。可见，一个工业产品的运动过程可图示如下：

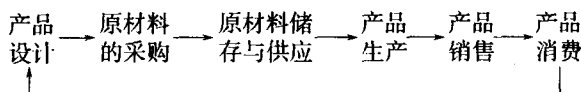


图4 工业产品的运动过程

统计信息作为一种统计产品，与物质产品的运动过程一样，也是一个由生产到消费的过程，全过程必须顺序经过以下六个环节：

**设计** 这是统计信息运动的开始，在某种程度上，也可以理解为准备阶段。这一阶段的主要任务是研究确定为了满足信宿的需要，搜集信源哪些方面属性的信息，以及属性的名称、信息载体的形式、信息的计量方法等。这是统计信息产品的设计阶段。

**搜集** 即根据指标设计的要求，对信源发出的信息进行量测和提取。这是为生产统计信息组织“原材料”的阶段。

**整理** 就是对所获得的信息进行确认，并条理地、系统地、科学地存储起来，为生产开发和利用者及时提供其所需的统计信息。这是统计信息产品生产过程中的原材料存储和供应阶段。

**开发** 就是以原始统计信息为对象，运用一系列科学的统计方法揭示出蕴含在原始统计信息中的实质性信息、预测性信息以及决策支持性信息。即统计信息产品的生产过程。这里需要指出的是，统计信息开发，除了统计主体的开发活动而外，也包括宿体的信息开发活动。

**提供** 运用各种各样的方式向全社会提供信息。这是统计产品由生产到消费的过程。没有这一过程，统计信息运动的一切消费都将是无益的。

**使用** 即信息使用者将所获得的统计信息用于认识世界和

改造世界的活动。在这一过程中,统计信息生产者,既要向用户作深入细致地解释、分析,帮助用户正确理解和掌握信息,并负责指导使用;又要负责研究、解决用户在使用中出现的问题,全面搜集用户对统计信息质量的评价意见和希望,即做好售后服务和跟踪反馈。在信息使用的同时,又生产出新的信息,于是又开始了一次新的信息运动。

综上所述,统计信息运动过程可图示如下:

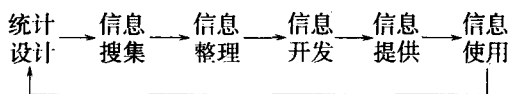


图5 统计信息的运动过程

### 1.1.3 统计信息系统

上面是就统计信息系统的运行过程所作的纵向透视。现在我们对统计信息系统横向剖析。要顺利地完统计信息产品的运动过程,一个科学高效的统计信息系统应由统计信息设计、统计信息搜集、统计信息存储和统计信息开发四个子系统构成。

#### 一、统计信息设计子系统

它是一个综合协调器。其输入是信宿对统计信息的需要,以及对所使用的统计信息的评价和要求;输出是设计方案。即:

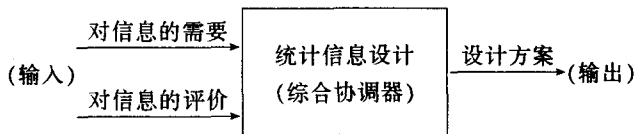


图6 统计信息设计子系统

该子系统的功能在于使统计信息系统的各个子系统协调运行,各个环节顺利进行。这一子系统的主要模块是统计指标数据库。

## 二、统计信息搜集子系统

统计信息搜集子系统是一个“采集器”。这个采集器有两个任务。一是提取信源所生产的信息；二是搜索信宿对所使用信息的评价和希望。所以，该子系统的输入有设计方案、信源产生的信息以及信宿对使用信息的评价；而输出为载荷所采集信息的信号。即：

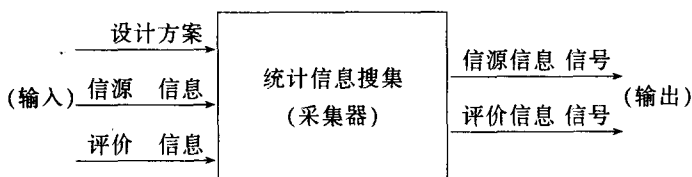
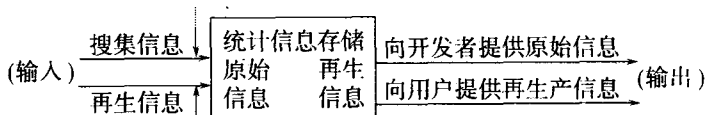


图7 统计信息搜集子系统

可见，该子系统的功能是提取、计量信息，跟踪、反馈信息。其主要内容是统计信息搜集方法库。

## 三、统计信息存储子系统。

显然它是一个存储器。这个存储器由两个模块组成：一是原始(一次)信息，二是再生信息。所以其输入有两个方面：统计信息搜集子系统的输出和开发子系统的输出。它将形成两个结果：原始信息库和再生信息库。由于该子系统的功能仅在于存储，因而其输入、输出基本一致。即：



注：——→表示噪声

图8 统计信息存储子系统

该子系统的功能在于确认、存储、提供原始信息(统计整理)和存储、提供再生信息。简而言之，即确认、存储和提供信息。

在这一子系统中,除统计信息库外,还必须包括统计信息质量监测控制模块、统计信息检索模块、统计图表生成模块及统计信息安全保密模块等内容。

#### 四、统计信息开发子系统。

它是一个生成器。即根据信宿的要求,以原始信息为材料生产出再生信息。即:

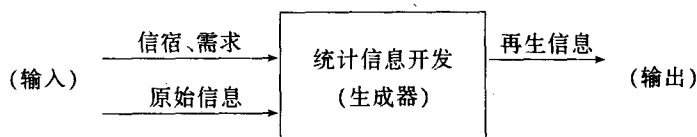


图9 统计信息开发子系统

该子系统的功能就是生产深层次的统计信息。其主要内容包括统计分析方法库和统计模型库。

综上所述,统计信息系统的结构和运行可图示如下:

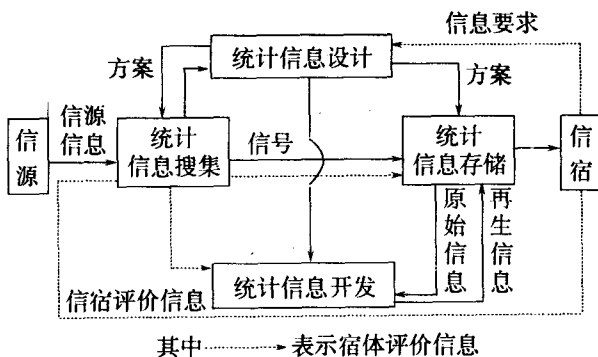


图10 统计信息系统的结构及运行过程

统计信息系统的结构及其行为,决定了该系统的整体功能就是为全社会生产和提供有价值的统计信息。

可见,在统计实践中操作的对象既不是物质也不是概念,而

是信息;操作工具既不是物质手段也不是思维工具,而是信息获取、信息加工手段;操作结果既不是新的物质产品也不是新的概念,而是新的信息。所以统计实践既不是在物质世界进行的物理操作,也不是在精神世界进行的思维操作,而是在信息世界进行的信息操作。在这个信息操作过程中,又有效、合理地运用了实践操作和思维操作。

## 1.2 统计信息系统:一种认识运动

通过上述分析,可以发现统计活动旨在认识事物、过程、关系、规律,所以统计活动属于认识活动;统计认识又是借助于信息这个中介体实现的,因而统计认识既不是一种物质运动,也不是一种思维运动,而是一种信息运动。信息世界是物质世界与精神世界的中介,而信息运动是物质运动与思维运动的综合。

### 1.2.1 信息是统计认识的中介体

统计认识与其它科学认识不同,统计认识所依据的第一手感性材料既不是客体的物理粒子,也不是业已形成的一系列概念,而是载荷着客体历史行为、发展变化轨迹、与其它客体相互作用的方式和程度、以及现在的状态和特征等信息的客观记录。因此,在统计认识活动中,为了能够在主体中反映客体,必须从客体出发把信息从客体传递到主体的物质流。这样,在统计认识中所观察的相互作用不是从“动力”角度,而是从“信息”角度去认识的。在对认识过程进行的分析中,应当将客体和主体之间传递信息的物质流——信息媒介作为一个重要成分包括进去。

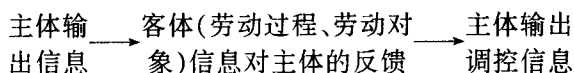
信息媒介实现着客体与主体特有的物质联系,同时又不以从客体向主体转移物质和能量为前提。这就保证了“客体——主体”的相互作用具有信息性质。根据这种观点,统计认识过程

是“从认识客体向主体转移信息”的过程。即主体对客体的信息进行搜集、辨识、储存、评价、加工、改造、选择、建构,并创造出新的信息的过程。

在整个认识过程中,统计认识主体既不对客体进行物理解剖,也不进行化学分解,而是对其历史行为和现在的状态进行观测。当然,这种观测也不是完全被动的,主体可以能动地作用于信息源(客体),可以创造一定的环境,使客体在特定条件下表现出自己的特征。如实验、试验、重复观察等。

### 1.2.2 统计认识是客体信息的一系列变换

劳动本身要完成的也只不过是把主体认识中的目的性信息通过计划信息的实践转化为客体的结构信息。劳动活动本身是一个不断进行着的:



这样一个信息循环运动过程。因此,从信息流这个层面来看,整个人类活动可以抽象为信息在客体、主体和宿体之间的循环往复:

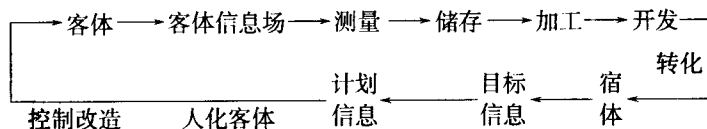


图 11 人类活动的信息流

统计作为人类活动的一种形式,旨在搜集、存储、开发信息产品,属于人类认识活动的范畴。统计认识是统计主体通过对客体系统的一系列建构,使客体信息作相应的传递、转换来实现的。统计主体在认识过程中所作的一系列系统建构可以概括为:



客体系统→概念系统→原始信息系统→统计模型系统→决策系统→控制系统

图 12 统计信息建构过程

概念系统的建构使客体系统转化为指标系统,使客体系统的结构转化为指标结构;原始信息系统建构使客体系统转化为信息系统,使客体的结构映射到数据系统之中,从而使客体的特征转化为数据的特征;模型系统建构,使客体的特征突现出来,转化为看得见、摸得着,便于认识、便于操作;决策系统建构将系统的运行规律、发展趋势与环境的影响以及宿体的愿望和目标有效地结合起来,得到整体优化的运行方案,并将客观信息转化为计划信息;控制系统建构使计划信息转化为指令信息,实现宿体科学地、客观地、合目的地控制、改造客体的目的。

统计认识方法是一种信息处理方法。它从客体系统的信息流出发,把系统看作是借助于信息输入、传递、加工、处理、藕合而实现其有目的性运动的一种研究方法。在分析、处理问题时以信息论为基础,完全撇开客体系统的具体形态和具体构成,把系统有目的运动抽象为一个信息变换过程:

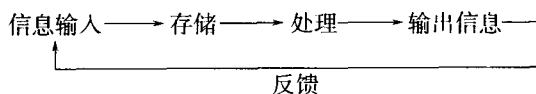


图 13 统计信息变换过程

信息变换过程是所有相互作用的事物之间普遍具有的一种现象。在人们意识之外存在着一个客观自在的信息世界,人们正是通过对客观信息的把握来认识自然的。统计信息搜集和储存的过程是统计认识主体把客观自在信息上升为主观自为信息的过程,而统计加工,统计分析是对自为信息的改造,是创造新信息的过程,统计决策是如何把目的信息转化为自在信息的计划信息。统计信息反馈则又是把自在信息输入人的观念,形成

新的自为信息,开始新的信息运动。

### 1.2.3 统计信息活动是一种认识活动

通过上述分析,我们发现统计活动系统的输入和输出都是信息,因而,统计活动是一种信息活动。但是,统计信息活动作为一种独特的形式与其它信息活动又有着显著的区别。

首先在于客体是统计信息活动的重要要素。它既是统计信息活动的出发点也是归宿点。而在其它信息系统中,客体是作为信源间接地出现的。仅仅是作为原因,而不是认识的目的。

其次,统计认识主体直接参与统计信息活动。统计认识主体必须从具有结构等级的客体中区分出结构本身。主体的这种活动建立在对客体 and 结构进行识别的能力上,建立在传递这个客体结构特征的信号之性能上。

再次,在统计认识活动中,信号(统计认识成果)对于主体和宿体来说,都具有“内部意义”。且这个“内部意义”是由作为信息源的客体决定的,是客体特征、规律的揭示、描述和反映,是客观认识的属性,并与客体保持着同一。这种属性构成了统计信息的语义方面。这不同于信息在控制论系统中的“内部意义”,在那里,所谓其意义就是指具有指令意义,带有操作的含义。统计信息也具有这一点。

统计认识成果——统计信息的“内部意义”之所以能与客体保持“同一”,是因为统计认识主体参与了从客体区分特定结构的的活动。统计成果和客体的联系问题是解释主体行为的关键。传统的认识论在考察客体——主体的关系时,往往侧重于主体的精神活动方面,而对客体与主体相互作用的机制本身注意不够。马克思主义认识论承认思维和存在具有同一性,并且“同一”于存在,思维是存在的反映。

最后,通讯实践中的信息是在对物质载体进行物理变换中保留下来的不变的东西——“结构”。结构与价值相似,看不到、

摸不着,然而却是完全客观的。

控制系统中的信息,除了结构之外,包含了对接收系统具有指令的性质,带有操作的含义。在这里只有具有指令含义的信号才是信息。

统计主体在统计信息活动中从客体(信源)的属性、关系、性质等所有多样性中区分出某些有限的结构(成为信息过程的起点),探索从客体获取信息和对之进行加工开发的各种方法,这一切对认识主体的活动提出了全新的和专门的要求。这种活动已经不能仅仅归结为消极的接收、存储、传递和加工信息,而且必须与客体本身,以及信息载体进行能动地相互作用,并且要以使用一系列专门的操作为前提。因此,在这个层次上,我们遇到的不仅是信息的采集、获取、传递、变换,而且遇到了可以称之为认识的活动。

## 1.3 统计内容:客体的数值规律性

### 1.3.1 统计从量的方面把握客体

自从质的概念和量的概念出现以后,人们就开始重视从量的方面认识客体的重要性。起初是认识客体的发展总量、构成、发展变化情况、一般水平、差异、产生的原因,然后对事物的静态规律(分布规律)、动态规律、内部构成要素之间关系,事物与他事物之间的联系,以及未来变化趋势和在控制某个或某些变量的情况下现象可能出现的状态。统计的这种引导人们的思维接近真理逐渐把握世界的过程,是辩证法的从简单到复杂,从抽象到具体这一抽象思维规律的具体体现。是逻辑的和历史的统一,是辩证法本身的精神——发展的原则和联系的原则的实现。

但是,统计作为人类认识世界的一种活动,其所要认识的内

容并不在于个别具体事物在特定时空条件下的状态,而在于认识对象总体的各种规律性。这些数值规律性既包括客体的数量特征和内部结构比例,也包括客体质量互变的数量界限和客体与其环境之间的数值依存关系;既包括客体的变化规律性和发展趋势,也包括客体运动最佳状态的数量特征和达到最佳状态的控制参量。不论哪一方面,其属性为总体一般而不是个别特殊,这是统计认识内容的实质。统计描述(依据总体全面资料进行认识)是如此,统计推断(依据样本资料进行认识)也是如此。例如,不论是对客体量的认识,还是质的把握,都是通过客体的量的特征和表现出来的数量规律性来实现的。

### 1.3.2 统计揭示客体的统计规律性

统计观念认为任何单个事物都具有其特殊性、偶然性和随机性,但是并非杂乱无章、无规律可循。各个个体都围绕某一“定值”上下波动,这个“定值”就是客体的一般性、必然性和趋势性。而且,客体的这种一般性、必然性和趋势性不是独立存在的,而是寄寓于许多个体的特殊性、偶然性、随机性之中。同时,个体的任一特征、属性都是可以借助于信息进行观察和测量的,这些观测的信息在足够多的条件下可以消除或削弱偶然性,显现出必然性,即统计规律。所以统计认识把握的是客体的统计规律性。统计规律是客体属性和客观规律的基本趋势,是客体属性的客观规律在一定观测条件下的显现,随着新信息的不断获得,统计规律不断逼近客体的真值。

统计规律具体表现为大量现象的规律性和单个客体行为的概率规律性。大量现象的规律性即总体的规律性。构成总体的每一个元素现实地存在于每个给定的时刻,并且这些总体可能的规律性只是从整体上表征这个总体行为的不同属性、趋势、特点。这里统计方法是研究大量现象的工具。某些单个客体行为的概率规律性是指在一定条件下,某个现象以一定的概率发生。

其中,概率是该客体在不同的外部条件下某种行为潜在可能性的数量上的估计。这一概率表现为客体的这一行为实现情况的相对数,而且这个数就是对它的度量。这里统计的集合不是客体实际存在的集合,而是反映单个客体行为的各种可能现实的测量或观测的总和。因此,统计方法在这里是研究单个客体行为的手段,这个手段具有原则上的不确定性(非单值决定论)的特征。量子力学是统计学的这种应用之最为典型的一例。

统计规律与物理规律不同。对于物理规律来说,只要给定初始条件和边界条件,客体任一瞬间的状态就可唯一地确定。它不包含偶然作用,因而也就不需要概率尺度,或者说,概率是百分之百,即百分之百的发生或百分之百的不发生。统计规律性则必须运用概率尺度,它把大量个体特征(微观量)的平均视作总体的性质(宏观量)。统计规律的论断不用确切的判断词“一定发生”、“一定处在”或“一定不发生”、“一定不处在”等,它只给出发生的概率即“依一定的概率发生”。这是由于在一定宏观条件下微观个体运动的各种可能性不被宏观条件所决定,各个个体具有一定特殊性、偶然性,因而,宏观量必然是微观量的统计平均值。所以,统计规律性是表明整个系统所有组成要素的行为,由此反映出该系统的客观性质。

统计规律的发现是对哲学认识论的完善和发展。它直接涉及到对因果律的认识。在这一点上,哲学界曾经出现过不同的认识。坚持机械决定论者只承认客观上只有一种形式的因果律,完全否认统计规律的客观存在性。另一方面,坚持不可知论者则利用统计规律性否定机械的决定论,宣称推翻了因果律,进而否定客观世界存在规律性和可知性。唯物辩证法认为,客观世界的因果形式是多种多样的,单一决定论只是因果联系的一种形式,而统计规律则是因果联系的另一种形式。大量的科学事实和人类实践充分说明了辩证唯物主义这一论断的正确性。

统计规律只是作为占优势的趋势而出现,作为在一个范围

内用不同方式表现的同一属性而出现。这种属性在每一个单独的情况下是不可能被观测到的,只有在普遍的、大量的、整体上的情况才可观测到。统计规律所记录的是必然性不以“纯粹形态”出现时的这样一些情况,可以说必然性是通过大量偶然性,也就是通过偏离、涨落、个别例外而为自己开辟道路的。偶然性——并不像人们早先设想的,是我们对原因无知的结果,它不是主体的属性,而是客体的属性。必然性——这只是在大量偶然事件中占统治地位的趋势,它就是统计规律的内容。在这里规律表示的是一定的联系、关系、趋势的必然性质,而统计的(或概率的)则意味着,这个必然性是通过它的永远“不可实现”,即通过大量偶然的偏离、变异、振动而“凸现”。

## 1.4 统计方法:认识数值规律性的法则

### 1.4.1 统计方法的总原则是大量观测基础上的求同存异

统计方法是对偶然现象中的必然性的定量描述。系统的开放性往往从系统外部引入某些偶然因素,这样,系统本身的完全必然性特征就有可能被打破。因而,在现实中,完全的确定性过程或纯粹的偶然性过程是不存在的,而都是不同程度地兼有两个方面,亦即承认必然性与偶然性是内在统一的,个体的差异性中蕴含着总体的同一性。统计方法就是运用科学的手段抽象掉各个个体的差异性,探求总体的同一性,并用差异性去标志同一性的内在质量。差异性统计产生和存在的前提,没有差异就没有统计;而同一性则是统计的目的,为了求得同一性才需要进行统计。统计求同存异的法则很好地体现了“亦此亦彼”,“既是偶然的,又是必然的”的辩证法思想,并用概率予以科学、客观、精确、清晰的描述。

偶然性与必然性在一定条件下可以互相转化。热力学第二

定律认为,系统内部的偶然性微扰(微涨落)间的协同相干作用可能引起系统整体结构的必然性的有序性变化;而一个具有有序结构的孤立系统,随着时间的推移则只可能向熵增的无序化方向变化,直到达到最大熵(随机性最大)的最无序状态为止。即必然性向偶然性转化。一个开放系统,随着负熵流的引入,则会使系统由偶然性向必然性转化,建构有序。概率统计对负熵的测定,就是对建构有序的描述。

人们欲认识客体某一方面的数值规律性,就必须选用与之相应的统计方法。只有这样,客体的某种数值规律性才能显现出来。若仅凭借人们的自然感官和自然脑力,那么认识则是很肤浅的,甚至是全然不可能的。在这个意义上说,统计方法是人类感官的自我延伸,统计方法的科学性和先进性直接决定着人们认识事物数值规律性的深度和精度。它把人的认识引入到人的自然感官不可能达到的更深层次,使人类的认识逐步向客观世界逼近。

描述统计方法致力于整理、分类和分析大量的原始材料。在关于所研究客体的本质的一定理论前提和清晰地理解研究目的和任务的基础上,对这些材料进行收集,进而进行统计加工,得出统计事实。所谓统计事实就是某种概括的经验陈述,这是对直接的实验资料、观察或测量结果的统计总结。因此,描述统计在本质上表现为由原始信息的原始材料而来的数学的表象方式或形成方法的多样性。统计学制定了一系列允许从原始材料获得这类统计事实的技术方法。

概率统计方法是“借此我们能够确定某事件在给定环境下发生的概率,对此环境我们一无所知,仅知在同一个环境下,该事件曾经成功了几次,又失败了几次”(The History of Statistics in the 17 and 18 Centuries, P363),即已知事件  $M$  在  $n = p + q$  次试验中成功了  $P$  次,失败了  $q$  次,求  $M$  的未知概率位于给定数  $a, b$  之间的概率,强调事件的概率确定而不能确知。

### 1.4.2 统计思维是一种系统思维

统计思维要求从传统的对象性思维方式转变为系统性思维方式。认识论上的“实物中心论”简单地说就是把研究对象仅仅作为一种个别的客体,而不是把作为客体的系统置于研究的中心。这种认识形式表现在科学思维上亦是就对象本身来认识的以分析、“解剖”为主导的所谓对象性思维方式。认识的“系统中心论”要求一切认识对象不仅它本身作为一个系统来认识,而且它还要作为某个更大系统的要素来认识。这种对系统客体的“主体”认识,是一种对研究对象进行整体性度量的系统思维方式。

统计认识在很多情况下是有效地应用了黑箱理论。一般来说,统计认识主体往往不知道认识客体的内部结构及其变化,至多也是知之甚少,而且也不便于或根本不可能进行“解剖”,因此,为了实现由不知到知,由知之甚少到知之甚多的转变,统计认识主体往往是通过观测客体系统的响应,根据客体系统的输出(响应)或输入和输出来进行探索和研究客体系统的结构和功能的。在这个意义上,完全可以说统计学是根据客体系统响应探求客体系统机理的方法论。由于系统响应事实上是客体系统内部所有要素综合作用并在一定环境约束下的结果,因而统计认识充分体现了系统论的整体观和开放观,这是统计研究的一大优点,也是统计认识比较接近客观、真实的主要原因之一。其它认识方法往往是就某一要素而研究某一要素,就某一系统而认识某一系统,忽略或没有充分重视各要素的整合作用和系统环境对系统的制约作用。

## 1.5 统计学:一门亚哲学

### 1.5.1 统计学的发展史是人类认识世界的历史

不考察统计的形成和发展过程,就无法弄清统计的本质。



所以应从历史的考察着手。统计学作为一门研究客体特征和规律的方法论学科,从其研究方法的发展过程来看,大致经历了以下几个阶段:

第一阶段就是《国势学》派的记述方法和比较方法。它不仅记述事实,而且试图通过比较来探索国家盛衰的因果关系。可见,就方法论意义而言,抽象掉其具体内容,其研究方法的功能在于认识现状,探索因果关系(规律)。应该明确指出的是,这一阶段的记述方法主要是文字记述,和表记、图示,而比较方法是形式逻辑的比较。此外还提出了根据“哲人的经验”和“历史的事实”进行推算的一般根据。例如,阿亨瓦尔(Gottfried Achenwall)在其代表作《近代欧洲主要国家的基本制度》一书中,在关于西班牙一章里,首先是西班牙的简史;接下来记述气候、地理位置、政治区划和物产,然后记述了根据纳税户推算出的乌塔兹(Uztariz)的居民人数;考察了造成人口稀少的原因,以及民族、科学、经济、军事等。

第二阶段是《政治算术》学派的大量观测方法、计量方法、分组方法、推算方法。如果说,记述方法是记述有无,侧重于对客体进行定性认识的话,那么大量观测方法、计量方法、推算方法则是计量“多少”,开始对客体进行定量认识,并抽象掉其个别差异,求取其同一性,即“去异求同”。

第三阶段是数理统计学的变量分析方法,旨在质与量的联系中认识总体异同的对立统一。在一定误差的基准上求其同一,并用“差异”标识“同一”。可见,统计学的发展史就是人类认识世界的历史,它反映着人类认识不断深化、准确、客观地认识和把握客观规律的过程。统计学所描述的是客体在一定时空条件下所处的状态,是客体本质的刻划,发展规律的揭示。所以,统计学每前进一步都标志着人类认识能力的提高并将对科学研究和管理科学起到巨大的推动作用。

### 1.5.2. 大量观测和实验是统计学产生和发展的基石

统计学不是一门单纯的理论学科,它是一种工具,一种语言,一种地地道道的普通方法论。它包含在科学活动的整体中,自身又浸透了科学活动的精神。

统计学研究的内容是如何获得信息(观测、试验、实验),提炼客体主要特征和规律性的信息。把握客体的变化发展趋势,研究调控决策方案。从科学研究来看,统计方法是从实践(现象)过渡到理论的中介。人们的认识是从实践开始的。统计认识在观测和实验的基础上,获取客体的状态、结构、内在联系、环境的适应性、运动规律等方面的信息。然后再经过科学研究及实践使之科学化、系统化、体系化形成新的理论。同样,在理论研究提出一种新的理论或假设的时候,往往要进行统计验证。从人类实践活动来看,一切人类有目的的活动都有具体的决策、行动计划、调控方案,统计不仅为制定决策和研究计划及调控方案提供科学依据,而且还是制定决策和调控方案的重要工具。在这个意义上说,统计活动不仅是一种认识过程,而且也是一种实践过程,包括实践活动计划的研究、实践过程的监测、调控以及实践结果的描述、检验及信息反馈等统计方法贯穿实践活动的始终。

统计科学是建立在大量观察和科学实验的基础之上的。一般说来,对于可控且宜于控制的现象来说,统计学强调进行大量重复实验(试验),而对于一些不可控或不适宜于控制操作的现象来说,通过大量观察来获得第一手资料——信息,在此基础上,寻找现象的规律性和创造获得这些规律的方法。可见,统计认识既不是单纯的逻辑演绎或推理,也不是只依靠经验。而是观察方法、实验(试验)方法与逻辑演绎和推理方法的有机结合。避免了重视思辩而轻视经验,只强调演绎而轻视归纳,只能够谈论而不能生育的不足。统计认识方法概括起来看,在搜集数据

获得信息(知识)的方法上,主要是大量观察法和实验与试验,而在整理经验材料、再生信息(即发展知识)的方法上,主要依靠归纳。所以统计学家既不是蚂蚁,也不是蜘蛛,而是蜜蜂。也就是说,统计学者既不能只会收集材料,也不能只吐丝造网制造知识体系,而是运用观察和实验(试验)的方法,从花园和田野里面的鲜花上采集材料,并运用统计推理方法根据这些材料创生出新的更深层次的信息。

### 1.5.3 统计学与实践

#### 一、实践是统计学产生和发展的源动力

统计学作为认识世界的科学理论,就其本质来说是唯物的、辩证的。它以客体的事实为认识的出发点,经过信息采集和加工,达到对客体状态的描述和结构及规律的揭示,并以与事实的相合程度作为衡量、检验、评价统计认识的标准。也就是说,认识的内容是客体的本质和规律。在这个意义上说,统计认识思维方法可以帮助人们养成尊重客观事实,实事求是,探索真理的辩证唯物主义精神。换句话说,统计认识自始至终贯穿着辩证唯物主义思想,是辩证唯物主义在认识中的具体实现。

统计学的发展,必须不断获得新的事实材料,仅仅在生产过程中以及对客体的直接观察是不够的,还有赖于各种仪器、设备的使用和不断发展,才能扩大认识的范围,开辟新的研究领域,获得更多的事实材料。然而观察、实验的仪器、设备的制造必须依赖于生产实践。如对天体的观察需要借助望远镜;对细胞的研究需要显微镜;对化学元素的研究,需要借助精密的天平和各种分析仪;对基本粒子的研究,需要借助高能加速器;对宇宙天文学的研究,需要借助航天工具以及大功率射电望远镜。采集和处理各种各样的大量信息需要借助多功能电子计算机和各种电子仪表,以及作出准确的判断(如气象预报、资料分布等),或者作为进一步研究的基础(如科研的情报资料),或者实现人类

对复杂大系统(如国民经济、交通运输等)的计划、控制和管理。统计的发展史表明,生产实践的发展,一方面为统计学开辟了更多新的认识领域,提供了更丰富的事实材料,另一方面为统计科学获取和加工、传输、整理信息提供了先进的技术手段。统计学的产生源于人类生产实践活动,统计学的发展必须依赖于认识世界和改造世界的发展,社会生产实践既是统计学产生的根本基础,也是统计学发展的真正动力。这是客观事实。

## 二、统计的能动作用

统计学发展的动力源于实践,但是,不能因此就认为统计学对社会实践总是处于被动的依赖地位。特别是在现代信息社会中,统计学对社会实践的作用更加突出、更加重要。统计学不仅是确定事实的记述统计学,而且是探索规律的分析统计学,还应是利用统计规律和统计事实预测未知的推断统计学,同时还是优选决策方案、反馈运行信息的控制统计学。统计学的使命在于引导人的思维去探索新成果,揭示自然和社会的奥秘,把握客观世界。概括地说,统计的基本功能在于提供统计信息产品,增强人类的认识能力和能动的调控能力,从而使可持续发展巨系统及其各个层次的子系统协调发展,实现良性循环。具体说来,表现在以下几个方面。

1. 描述职能:一切认识都是以确立事实,并以事实为依据开始的。统计学在占有大量观察、实验(试验)材料的基础上,对现象的状态、内部构成、分布特征等予以客观描述。实现对认识对象的初步整体认识,并为更深入的认识提供认识基础。

2. 发现职能:统计具有演绎推理的性质,在这一点上,可以认为统计只能解释、阐明已有的知识,而非获得新知识。但是统计作为不完全归纳推理和类比来说,根据已取得的统计信息,运用统计方法进行加工、整理,揭示出现象本身的内部联系和该现象与它现象之间的依存关系及其发展变化规律性,统计推理结论可以扩展到没有研究过的对象和现象,可以推出事物具有在

前提中尚未判明的某种属性与特性。将统计认识由描述状态推向发现规律。

统计学不提供确实无误的结论,它的结论常常是可以有错的。但它常常提供一种创新的结论,又常常提供足够合理的信念,在常识不能作选择的地方,促使我们接受或拒绝某个有争议的结论。更为重要的是,统计认识对认识深化提供一些重要的启示和进一步研究的问题、方向及思路。

从统计的观点接受某一个结论,并不是让你把它当作不变的真理来供奉和崇拜,而是让你去选择、去运用、去行动、去寻找问题和解决问题。

3.解释职能:有了对认识客体状态的描述和内在规律地发现,就可以对客体作出客观的解释,将统计认识转化为人类知识、社会财富,使人们能够据以正确理解客观现象,掌握客体的性质、特征和运动变化规律。

信息描述——即统计模型。客体是认识对象的物质方面,模型是对象的信息方面。一个合理的模型应保持与客体相同的结构。从理论上讲客体的物理结构与信息结构完全相同,是同一个事物的两个方面。但是,由于人们的认识不可能达到完全、客观、准确的程度,因而统计模型往往与客体的物理机制存在着一定的差异。这种差异就成了衡量、检验、评价模型优劣的依据。正是由于统计模型的这种同构性,使得人们对认识客体不仅有了量的概念和把握,而且可以远离庞大而又复杂的物质系统,通过模型进行模拟客体系统的结构和运行。为人们对外体进行更深入的认识和探索提供科学合理的手段,为实现对客体的最优控制及有效的改造提供有效途径。

4.实验职能:实验的概念往往是物理的操作,是人们在物质领域,通过控制物质要素来把握观察系统机理或是获得要素对系统的作用和影响。由于统计具有模拟功能,实现了系统在信息领域的同构,因而人们可以在信息领域进行实验,客观上称之

为试验。当然统计实验是利用统计模型,通过控制信息要素(要素的信息方面),来认识和把握认识对象的系统机理,获得要素对系统的作用和影响的信息。统计实验是在信息域进行的,信息可以离开其客体的物质态而存在。因而统计实验具有简单、易于控制、投入少、成本低、实验速度快等特点。

5. 预见职能:统计科学不仅是关于已经发生的客体状态和变化过程的记录、描述和认识,而且还提供了探索客体未来发展趋势的方法。统计学既可根据客体过去和现在的状态,按照客体自身发展的规律对其未来的趋势和可能达到的水平作出动态预测,也可以根据结构要素和主要影响因素,按照其内在联系和依存关系对客体的未来状态作出预测。统计学的预见职能使人类认识客观规律变事后为事前,变受动为主动,从而使人类认识客观世界的能动性具有了可能性。充分体现了统计学在人类认识中的特殊作用。

6. 控制职能:统计调控职能是通过将反馈信息与目标信息进行比较分析,并根据客体的机理,结合运行信息与目标信息的差量对系统的参数或(和)输入量进行修正,并转化为计划信息,作用于客体系统,从而保证系统按照预先设计的目标进行运转。

统计信息产品是客体系统物质运动的信息形态。虽然统计活动不直接提供物质产品,但是统计活动把物质的运行状况描绘出来,使寄寓于物质形态之中的、抽象的、隐藏的性质、规律、走势等属性,变为直观的、可以看得见、易于理解的信息,从而据以对所研究系统进行调控,使各系统按照人们的目标协调发展。简言之,统计信息可以提高人类认识能力、调控能力、促协能力。在这个意义上说,统计信息完全是人类生产力的构成要素之一,并且,其作用将随着科技的发展,越来越重要。

#### 1.5.4 统计观点作为哲学意识形态的地位

##### 一、因果观点的困境与统计观点的形成

在十九世纪以前,几乎所有哲学家都认为原因与结果之间的关联如同逻辑关联一样是必然的。宇宙的一切都受因果性制约,且这种因果性是必然的、有规律的。休谟认为“唯有因果关系使我们能够从某个事物或事件推论其它某个事物或事件:‘使我们由一对象的存在或作用确信它随后有、或以前有其它什么存在或作用,产生这种关系链的唯因果关系而已。’”(《西方哲学史》下卷(英)罗素著 商务印书馆 1976 年版 第 201 页)“因为甲,结果乙”这样形式的命题就是因果律。其中甲和乙是两类事件。

休谟把人类理性的一切对象分为两类:“观念的东西”和“实际的事情”。前者是数学研究的对象,因此,作为演绎逻辑的基本规律——矛盾律支配着“观念的东西”,而因果律支配着“实际事情”的关系。休谟指出:“关于实际事情的一切现象似乎都建立在因果关系上。只凭借这种关系,我们就可以超出我们记忆和感官的证据以外。”(《人类理解研究》,商务印书馆,1982 年,第 27 页)且认为因果关系的基础是经验。即同样的固定的前后相续出现,重复次数多了,便会使人们不由得从一个物象的出现立刻期待另一物象的存在。总之,因果观念是物象先后关系的重复作用于人们的效应,是人心由于习惯而生的信念。

十七、十八世纪牛顿力学把因果观点拥上哲学意识形态的地位。拉普拉斯认为“我们应当把宇宙的现状看作它先前状态的结果以及它的后继状态的原因”。在这种思维赖以进行的模式中,允许设想一个无限勤勉,又有无限精力的全知者,他能掌握某一时刻大千世界一切事物的状态,以前推千年,后测万事,甚至我们毋宁说,因果框架本身就是全知者的居所,因果性就是全知者的观点。人类的心智自然够不上全知的要求。但是,牛顿力学的基本定律借助于欧几里得几何学和微积分,可以对世界系统的过去和未来状态进行类似于全知者所作的那种分析,尽管还有无限多未知的细节留在后面,但人们曾经相信,它们全都能用同样的方式处理。牛顿力学使富于想象的人类心智

产生了一种幻觉,一种接近或达到全知者的奢望,或者反过来说也一样,全知者的地位正是人们对自己在牛顿力学中实际达到的观点的理想化。就这样,把因果观点拥上了哲学意识形态的地位。

然而,随着人类活动和科学研究的发展,物理学和其他学科碰上了一类新的客体:它们必须研究事件、事物、粒子、材料的大量集合、大量总和,以代替很容易被认为同质的少数客体。一句话,必须研究这样一种集合,其元素相互间并不是绝对同一的,它们在一定的范围内用不同的方式表现这些属性。这类大数现象更多地进入到像生物学、政治经济学、社会学等等这样一些学科和研究范围之中。为了研究这类大数现象,较早采用的方法似乎是不适用的,因此制定了新的研究、分类、概括和预言的方法。这些方法被称为统计方法。统计方法在颇大程度上是定量方法。用统计方法确定的属性、分布规律在开始时也就称为“被统计的规律性”,并且被看作是不精确的、近似的、粗略的东西,被看作是“次等的规律性”,被看作是我们关于“真实的”、“精确的”规律的知识之不完备的见证。在现代,对统计方法的本源和性质的理解由于对大量现象本身规律性的客观性质的更深刻的理解而放弃了。

例如,在气体分子运动力学这样的学科中,通过对大量微小元素的个别行为的统计假定,有可能得到一些刻画宏观总体特征的概念与定律,这往往是严格的因果观点所做不到的。最浅显的例子是“温度”概念,它是每个人凭感觉都知道的,但是单个分子并没有温度,温度是大量分子聚集起来后表现出的总体特征。且不说以严格因果观点写出无数条方程,一个一个地描述分子运动实际上不可行;即便真正做到了,也得不出与人们的经验感觉所对应的关于“温度”的理论概念。

因果框架容纳不下新的认识领域,如微观世界。当牛顿力学的框架难以支持自然界的丰富内涵之际,统计观点帮了它一



忙。借助于大量分子行为的统计假设,温度重新可以在牛顿力学原有基本概念上定义。热力学的全部定律也都可以用这种方式重新得出,其中最著名的就是波尔兹曼对热力学第二定律(即熵增加定律)的概率解释。

在这一阶段,统计观点成了因果观点不可或缺的补充,没有统计观点的帮助,严格的因果观点越来越显得并不能洞察一切。正是在这个意义上,统计观点显示了自己的力量和独立性。另一方面,统计力学中的概率观点毕竟仍然是与因果决定论的思维框架相协调的,是牛顿式建筑的和谐与自然的延伸,对大量元素行为的统计假定仍然使用基本的牛顿术语和概念,仍然以承认每个个别元素服从因果律为前提。

在量子领域,人们曾经作过巨大的努力,试图通过“折射”从因果观点看清这一领域,但总得不到所期待的图象,硬把牛顿经典物理学的因果范畴加给这一领域的微观客体,举出了本质上的不确定性和含糊性。但是从概率观点出发,却能对量子现象作出不含糊的描述,表述为严格的统计性规律。虽然不能因果地确定量子过程的观测结果,但能确定出某种结果的概率。换句话说,量子现象服从确定的统计性规律。

在社会领域里,大量社会实践表明现象与现象之间的关联更多地表现为相关关系,而并非是“完全因果关系”。更为严重的是,几乎所有现象都不遵从“因果单值决定论”。这种相关性只有在大量观测的基础上才现显出一般趋势性。

这时,统计观点已经不再是原则上可以放弃的权宜之计或可弥补的知识缺陷,也不再是虽不可放弃却只能起辅助作用的一根拐杖,在量子领域,统计观点成了整理和解释观测事实的唯一可能的观点,成了与因果观点不相容的第一性的基本的范畴,形成了一种新的意识形态,和因果观点一样,拥有了作为哲学意识形态的地位。

在因果观点中包含着一条甚至拉普拉斯都没有明确陈述出

来的前提,即认识主体与认识客体之间存在着截然分明的界限,假定认识主体独立于认识客体之外,与客体毫不相干,主体对客体的观测丝毫也不会干扰和影响客体的状态。事实上,在人类认识活动中,不论是在自然领域,还是在社会领域、思维领域;也不论是宏观领域,还是微观领域,不可避免地包含着“观测者的参预”,并且认识主体发挥着重要的能动作用。认识结果所描述的是由观测者参预而生的现象。主体的参预并充分发挥其能动性,成了人类认识的必要条件。

统计观点不仅在因果律之外的领域行之有效,而且从系统的观点来看,由于系统内部各要素之间的相互影响和相互作用以及系统与其环境之间的复杂关联,使得本来是因果关系的现象也往往表现出不确定性来。因此,常常需要用统计观点探索和研究某些现象之间的因果律。

统计认识既不是演绎逻辑,也不是经验主义的“习惯”或“惯性”,而是客体特征的信息检测。统计推断不是“枚举归纳推理”,按照休谟的思路,因果性被归结为人们对“实际的事情”之间牢固、恒定、无例外的前后关系的主观信念,其根源存乎人心从以往这种前后关系的多次重复而养成的“惯性”,以致一看到前一事件出现,就必然会期待后一事件也出现。“枚举归纳”并不具有客观的逻辑上和经验上的依据。统计归纳一开始就假定过去与未来,已知对未知的关系是概率性的。

## 二、统计方法的普适性决定了统计学是一门亚哲学

统计方法适用于一切科学。现实世界中任何一种事物及其运动形式都有其质的规定性,又有其量的规定性,而且质一般要通过一定的量来表示。统计学就是研究量、量的关系和量的变化从而揭示研究对象的运动规律的科学。只要是对于量的研究,就可以运用统计方法。运用统计方法可以对客体进行更为精确的定量描述,获得前所未知的参数和关系的信息,可以判定一些数与界限清晰的概念,不仅可以预测未来,而且还可能导致

思想和概念的形成。由于统计学所研究的量及其关系普遍存在于各种物质系统及其运动形式之中,因此,各个领域一切科学技术研究原则上都可以用统计学来解决有关问题。其应用广度是空前的。

另一方面,系统的普遍适用性也决定了统计方法的普适性。统计认识以信息系统为中介研究客体系统,虽然客体系统千差万别,但都可抽象为信息运动系统,因此都可以运用统计方法进行信息加工处理,达到对客体特征和规律性的把握。

信息的哲学意义决定了统计学的亚哲学地位。广义统计信息从共时态来考察,有社会信息、自然信息和思维信息。这些客体信息形成了统计方法论研究的客体;由于不同的客体信息具有各自特殊的变化规律,因而形成了相应的应用统计学科体系:社会统计学、自然统计学以及思维统计学。从历时态来考察有历史信息、实时信息和预测信息,因而有动态统计学、静态统计学和统计预测学。从统计信息的可测度来看,有确定性信息、随机性信息和模糊信息,相应的有描述统计学、推断统计学和模糊统计学。从统计信息的效能来考察,有监测预警信息、控制(咨询、监督)信息和决策信息,同样,有预警统计学、统计控制、统计决策。所以说统计学就是研究信息的测量(搜集)、存储、传递、处理、提供和利用的方法论学科;统计思想的核心就是通过大量观测研究,消除特殊性和随机性,认识总体的一般性和必然性;统计思维和统计方法具有普适性,它适用于认识自然、社会及思维等各个领域,是人们认识世界的重要手段,因而在整个人类知识体系中处于次哲学的层次,所以我们称之为亚哲学。说它是一门“哲学”,是因为它和哲学一样,都是认识方法论,具有普遍适用性;它告诉人们如何认识世界;其实践行为属于认识世界的范畴。说它是“亚”哲学是因为哲学为人们提供认识世界的立场、观点和一般性原则,即世界观,而统计学则为人们提供从数量方面认识世界的具体方法。

### 三、统计学与哲学

辩证唯物主义认识论和统计学都是认识论。前者是人类知识的最高总结和概括,揭示世界的一般规律和人类的总看法,对各部门具体科学都有普遍意义;后者则是研究人们如何从量的方面认识和把握客体,为各种科学提供一种以信息为媒介的认识方法,对客体的状态、结构和规律给出量的描述,使客体成为可以感知的东西,使人们的认识具有量的概念和界限,从而更清晰、更深入、更具体、更精确。

哲学是关于世界观和方法论的学说,它研究自然、社会和思维的最一般的规律。它和统计学是一般和个别、共性和个性的关系。哲学对统计学起指导作用,为统计科学研究和统计工作提供一般指导原则和思维方法。统计学是哲学一般认识方法的具体化。因为人们从事一切统计认识活动,毫无例外地存在着统计认识主体与统计认识客体及其环境(自然的、社会的)的关系问题。如何处理认识主体与客体及环境的关系,首先涉及到的就是哲学的指导思想问题。是从客观出发,还是从主观出发?客体是孤立的,还是与它事物处于一共同体之中?客体是静止的、绝对的,还是发展变化的、相对的?客体的本质及规律是可知的,还是不可知的?主体的认识结构是先天固有的,还是后天不断架构的?统计认识是从客体本身着手,还是从先验论出发?统计认识结果的检验标准是客体的运行,还是主体的臆断等等?这一系列问题都与哲学思想密切相关,只有以正确的哲学理论为指导,才能正确处理和指导上述各种关系,使统计认识不断向深度和广度进军。

其次,统计学是一门认识方法论,统计活动是一种认识活动,是要研究探索和发现认识客体本质及其规律性的方法。这就必须通过辩证思维,由表及里、去粗取精、去伪存真,才有可能做到。因为任何一个被认识客体都不是孤立存在的,也不是直接显现其本质的。总是和其它客体相互影响、相互依存、相互作

用错综复杂地共处一个共同体之中,本质寄寓于表象之中,必然存在于偶然之中。如何才能抓住客体的本质属性;如何进行观察或实验,如何采集信息,用什么样的方法进行数据处理,怎样消除偶然显现必然,以达到认识一般,又怎样消除共性显示差异以区分不同类、属、别或者权衡构成要素的不同作用。共性的发现和差异的确立对于认识活动来说,都十分重要。同质性显现共性,差异性显现不同类别。共性在一定范围确定同质总体(类、属、别)的范围。差异超出这个一定范围,就表现为一个新质总体(类、属、别),这个界限就是不同类事物的界限,就是客体质变的界限,就是哲学上的度。

统计研究和实践工作者,必须坚持世界统一于物质,物质是不依赖于人的意识而客观存在的实在、世界是可知的,一切事物都存在着普遍联系和运动演化等辩证唯物主义基本原理和原则。

#### 四、统计学与其他学科的关系

在科学体系的发展过程中,统计学与其它科学互相依存、互相促进。统计科学区别于其它科学的显著特点在于它的普遍实用性。研究对象也不是某一个特殊领域,而是各个领域的信息。

1. 统计学作为一种科学实验方法,是其它科学理论赖以建立的基础。

首先,统计学为一切科学理论积累事实材料,为科学理论的建立提供根据和来源。科学理论是关于现象的本质及其运动规律的系统知识,一般是从大量的事实材料中抽象和概括出来的事物的本质联系,是从感性认识中提升出来的理性认识。然而,社会实践所提供的大量材料是比较粗糙,比较零碎的,无法深入地、系统地揭示客观事物的本质和规律。统计学在占有大量观察、实验材料的基础上,运用统计方法和统计模型,对现象的状态、内部构成、分布特征等予以客观描述,从而为其它科学研究进行理论概括提供比较可靠的事实材料,为更深入地探明客观

事物的内在本质及其发展变化规律提供认识基础。

其次,统计学丰富完善原有的其它科学理论,并促进其不断发展。任何科学理论都经历了一个由不太完善到完善、由不太成熟到成熟的发展过程。在这个过程中,统计学起着补充、修订的作用,这是不容忽略的问题。同时,由于统计学的模拟和实验职能,人们可以利用统计模型来认识和把握研究对象的内在机理,但有时也会出现一些意想不到的新问题、新现象,需要进行新的理论说明与概括,这就为其它科学理论的发展开辟了新的领域,为其提供了发展的动力和源泉。因此,随着统计认识水平的不断提高,认识手段的不断改进,统计学对发展其它科学理论的作用将越来越重要。

第三,统计实验是验证科学假说使之成为科学理论的最有效的手段。人们在认识世界、改造世界的社会实践中提出的科学假说往往不能直接得到验证,而必须借助于科学的实验方法来验证。统计学通过大量观察和科学实验,并结合运用逻辑演绎和推理的方法,从偶然中发现必然,反映事物普遍的、一般的、本质的规律,从而成为检验真理的重要手段。

## 2. 统计学需要其它科学理论的指导

统计学在引导人的思维去探索和把握客观世界过程中,往往针对某一具体系统进行观察与实验。在对该系统的状态、内部构成、分布特征及发展变化规律进行客观描述和分析时,需要以该系统为研究对象的相关科学理论作为其理论指导。这些相关科学理论,是在一些基本概念、原理基础上通过逻辑推理而形成的逻辑知识体系。运用统计方法对某具体领域进行观察与认识时,可以根据相应领域的科学理论所形成的逻辑知识体系,作进一步的统计推断。对具体事物的质的界定,指标的含义、口径与范围,以及现象变动的原因说明等,必须以所研究领域的相应科学理论为指导,以其所形成的逻辑知识体系为依据作进一步的统计推断。尤其是统计学所应用的一些公式、模型、以及参数

的估计与检验方法等,一定要符合客观事物的内在规律,并为相应领域的科学理论所证明。无论研究领域是社会领域、还是自然领域,皆概不例外。同时,统计认识是一个由感性认识上升到理性认识的过程,统计学不仅仅具有模拟和实验职能,更重要的是在此基础上对统计实验的结果进行分析、判断、概括和总结,这些也必须结合运用相关的科学理论才能得出正确的结论,实现对事物的质的更高层次的认识。

## 2 统计认识论

### 2.1 统计认识一般

统计认识作为一种信息活动旨在认识客体、观念地把握客体。这种认识不是机械地摹写、盲目地跟随客体,而是创造性地反映客体。统计认识以关于事物、过程和其它形式的客观实在在人的意识之外并且不依赖于人的意识而存在这一唯物主义观点为出发点。但统计认识同客体本身并非一回事,统计认识属于信息的范畴,因此,统计认识与客体相对立。但是,需要强调的是对立本身是一种联系,而不是割裂,也就是说,统计认识同客体既相互对立,又相互一致,构成对立面的统一。它们是统一的,是因为统计认识是客体的再现;但它们又是对立的,因为统计认识与客体的关系表现为认识与被认识的关系。统计认识作为客体的再现不是以客体的物质形式再现。客体的物质再现(如物理模型)与客体是同一类的,其差异表现为同一共



性范围内的差异。而统计认识是对客体属性的把握,但并不具有客体的物理上的、化学上的、功能上的属性。

这里强调统计认识创造性地反映客体是指统计认识所把握的客体,不仅是在一定对应条件下存在的那种样子的客体,而且又是凭借主体的理性推断在客体各种可能变化形式,各种可能出现状态以及各种发展趋势中的客体。统计认识再现客体,既不是以某种物质形式复制客体,也不是创造与客体相同的东西,而是统计认识主体在客体的运动中观念地把握客体的特性和规律性。在整个统计认识活动中,统计认识主体并不改变客体,仅仅是再现客体的规律性。

### 2.1.1 统计认识运动

统计认识是人类活动的一种特殊形式,是一种既不同于物质运动,也不同于思维运动的一种信息运动。包括信息的提取、传输、加工、储存、开发、提供和使用指导。客体本身的运动同其物质特性的变化有关。物质的一种运动形式能转化为另一种运动形式,或者一个客体可以转化为具有新质(新的属性)的另一个客体;在物质活动中,与人直接打交道的是具体的物质。思维运动是一种精神的运动,是观念的运动,是思想、观念、意识的生产;思维的运动导致一种认识映像为另一种认识映像所代替,引起从不知到知,从对象客体表象的和片面的知识向深刻的和全面的知识的转化;在思维运动中,与人直接打交道的是早先形成的概念和业已达到的科学认知体系。信息运动既不是作为客观实在的客体的运动,也不是作为主观存在的观念运动,而是作为客观存在而不实在的信息的运动。信息运动是由一种特质的信息向另一种特质信息的转化,这种转化既可以是同一层次不同方面的横向扩展,也可以是向更深刻、更真切的纵向深入。信息运动又是以客体运动为基础的,和客体运动规律是一致的。信息是客体的信息,没有客体的运动就没有信息,更无信息运动可

言。

信息既是客体的一种属性,但是信息又具有相对独立性。它可以从客体的运动中分离出来,记录下来,并按照其自身的规律进行运动。例如,某地区粮食产量 10 亿公斤,一旦将这个数量信息记录下来,它就远离了粮食本身独立存在。同时,既可存储、传输,也可同有关信息进行综合,加工生产出新的信息,如人均粮食产量、亩产量增减幅度等。思维活动是观念的运动。观念是客观的主观反映;但往往人们不是通过直接与客体打交道,通过直接感知客体而形成观念,而往往是通过对客体的信息研究,来形成观念。例如,要了解我国钢铁生产情况,不需要将全年生产的钢铁集中到一起进行研究,而是通过钢铁生产的信息来达到认识的目的。再如掌握某个人的心脏功能,不是通过对现实的心脏进行解剖,而是通过对心电图的信息进行分析来实现的。这就是物质运动、信息运动和思维运动三者既相对独立,又紧密地统一于一体的辩证关系。物质运动是基础,信息运动是中介,思维运动是通过信息运动实现对物质运动的认识和把握。统计认识从目的性来看,属于思维的范畴,但从对象来看,处理的和运动的是信息而不是观念。也就是说在统计认识运动中(生产活动中),生产的材料和加工出来的产品都是信息,而不是物质和观念。因此,可以说统计认识是认识主体根据客体的运动规律,运用思维运动方式的一种信息运动,是物质运动、思维运动和信息运动的统一。

### 2.1.2 统计认识的基本属性

统计认识是统计主体在信息活动中,通过统计思维创造性地反映客体的过程。这体现了统计认识的客观性,人的主观能动性和认识的发展性。

#### 一、统计认识的客观性

首先,表现在统计认识是以关于事物、过程和其它形式的客

观实在在人的意识之外,并且不依赖人的意识而存在这一唯物主义论点为出发点。同时,规定了统计认识不是记号或符号,而是客体的属性和规律。记号或符号是形式,数理推演是手段,客体属性是内容。因此,统计认识从其内容来看,毫无疑问是客观的,是与客体相一致的,从其所用的认识工具——统计方法来看,方法本身也是客观的,不依赖于某一个统计认识主体的个人意志为转移。例如,同一资料的均值与方差不论谁来计算都是一样的,否则,就是有错误的操作。

其次,认识主体选择统计认识方法、方式、工具不是任意的、随心所欲的,而是由研究对象的特性和研究目的决定的。统计认识只能根据客体选择适合的手段,这说明统计认识方法和认识工具不仅没有改变认识的客观性,反而要受客体的制约。进一步讲,认识手段与客体相互作用所形成的状态和关系也是客观的,绝不会使客体依赖于主体而丧失客观性。

再次,统计认识主体的认知结构对主体的认识起着一种选择、定向、整合、建构的作用,但是它也不能改变认识的客观性。首先统计认识主体的认知结构不是先验的,而是实践的內化,是人类认识长期发展的凝结物。它是由心理因素、知识经验和概念系统构成的一种认识框架,其核心是范畴、概念系统对信息的规范、整合的作用。指标体系、公式体系是经过历代千百万次的反复实践而固定下来的,并将随着整个社会实践的发展而发展。其次认知结构为客体所规定和规范,认知结构的作用必须以客体信息存在为基础,如果没有客体信息输入,认知结构就没有加工整理的信息素材。采集信息实质上是从客体的机理中析取出有关客体某些特征、规律和运动过程的信息,并以可感知的符号记录下来。就是加工、整理,也不是任意的,必须遵循统计认识的法则和规律。

总之,在统计认识中,统计认识主体与客体的关系并不完全和意识同物质之间的关系相一致。客体并不是简单地作为不依

赖于意识而存在的客观实在的物质,而是全部这种包括在主体之认识的、实践的、变革的活动中的客观实在性。主体也不是简单地作为高度组织起来的物质属性的意识和心理,而是社会的人。主体实际上并不改变客体,而仅仅反映客体、认识客体的规律性。主体对客体的理论上的关系是在实践、观测、实验关系的基础上产生和存在的。统计认识作为主体的信息活动,不是把主体与客体割裂开来、对立起来,而是把二者结合起来、统一起来,这种结合和统一具体表现为统计认识是主体对客体的主观反映。

## 二、统计认识的主观能动性

统计认识主体的主观能动性主要表现在统计认识始终属于作为统计认识主体的人,是人在操作,个人的活动与信息活动相结合的运动过程。不论是采集信息,还是选择方法,或者是确认,都依赖于认识主体的认知体系、认识能力和认识行为,同一客体不同的统计认识主体可能会得出不同的认识结果。因此,主体的主观因素都或多或少地参与活动并发挥作用;再者,统计认识成果(统计认识活动的产品)不是客体本身而是对客体性质、规律的观念把握,是统计认识主体建构的产物。统计认识的结果不是认识客体和客体所具有的各种属性本身,而只是客体的相应的信息并表述为主体(观念)映像;统计认识过程总是同信息打交道而不直接同客体本身打交道;统计认识对客体的刻划、揭示、反映的完备程度、真实程度和深刻程度是不同的,统计认识在揭示客体时难免有片面性、失真性、甚至会出现歪曲。说明统计认识不是简单的记录和机械的摹写,不是照镜子,不是“照像”,也不仅是“透视”,而是属性的刻划和规律(机理)的揭示。(是照像、透视、解剖、脑血流图等各种手段的综合)。

## 三、统计认识的发展性

统计认识的发展性表明统计认识活动以辩证法为指导。一方面统计认识不是一项活动而是一个过程,统计认识不能一蹴

而就,而是逐渐接近。另一方面统计认识是一个实践→认识→再实践→再认识的循环。决定了统计认识的内容不是静止的、僵死的,而是随着认识的循环往复不断利用反馈信息和新信息进行校正、深化、抽象。从不知到知之,从低级到高级,从浅到深螺旋式上升的辩证发展过程。同时统计认识方法体系也是一个不断丰富和完善的开放系统。

此外,统计认识还具有超前性的特点。统计认识既研究客体过去的运动轨迹、动态结构及运动规律性,也认识现时刻所处的状态和未来的发展趋势,是超前的认识论,统计思维方式具有超前和创造的特征。统计认识不是客体状态的机械反映、被动反映,而要揭示客体的本质、结构和规律性,要根据其本质和规律性推断未知现象,预见未来时刻的发展趋向,也就是观念地创造未来客体的状态。从统计认识的程序上来看,统计认识往往从描述状态着手、从已知出发,但是统计认识的思维指向是未知的东西,是未来的状态,着眼点在于推断和预测。统计推断和统计预测都是超越现实,跃出已知的时空范围,利用现有的统计认知结构和认识手段,根据已掌握的客体信息对其未知状况或未来状态预先作出推知或判断。

#### 四、对统计认识的批判

1. 由于统计认识存在着主观性和客观性两个方面,统计认识的客观性可以保证统计认识减少主体的主观随意性,使认识结果与客体一致,但是也不能过分地夸大客观方面,更不能使之绝对化,否则就变成了机械地摹写,而不是能动的反映。事实上,完全抛开主体的主观能动性,就不存在认识。

2. 另一方面,统计认识的主观方面为统计认识主体充分发挥主观能动性提供了条件和可能,使统计认识成为一种创造性的成果,是人类的宝贵财富,但是,也不能过分夸大,也不能绝对化。否则,统计认识就成了随心所欲,就不是对客体的反映,而成了对宿体的满足。

3. 统计认识受人的社会历史条件以及科学水平和技术手段的制约,因而,就一次统计认识活动来说,统计认识客体的程度是有限的,即使是一次非常完美的统计认识活动,也不可能做到对客体的彻底把握。

4. 统计认识在反映客体时,难免有片面性,难免有把握不当的情况。

5. 统计认识结果受多种因素的影响,如客体的复杂程度、客体的可测量性、客体与环境的关系、客体内部结构与外部影响因素、主体的知识结构、主体选择方法的正确性、统计加工运算的准确性、指标设计的科学性等,因此,统计认识不可能全面考察各个方面,不可避免地会遗漏某些重要方面,或对某些重要方面重视不够,而影响认识结果。

6. 统计认识是对客体的反映,但是如违背了统计认识规则,统计认识就成了对客体的歪曲。

7. 统计方法是已有一定知识体系的产物,存在着滞后性。

8. 统计认识可信而不可靠。统计认识方法是科学的、统计认识思维是严密的、统计认识内容是全面、客观、精确的。所以是可信赖的。但是,正是由于统计认识的是客体的总体的、一般的、规律的、概性的。所以它抽象掉了局部,抽象掉了个别,抽象掉了具体,抽象掉了形式,抽象掉了表现,因此,统计认识结果和某一具体表现往往存在着量的差异。这个差异量是偶然与必然的差异,是个体与总体的差异,是个别与一般的差异。应该说是内在的、必然的,而不是偶然的。所以说统计认识是不可靠的。如统计认识研究某一地区男性的平均寿命为 72 岁,但丝毫不意味着张三的寿命就不会高于或低于 72 岁;说某一时期某地区的零售物价总水平上涨 2%,但是就电视机或任何一种其它商品的价格都可能没有变化,或涨跌幅大于、小于 2%。等等。

我们对统计认识给予批判,不是为了否定统计认识,而是为了提醒我们辩证理解、恰当运用统计认识成果。并激发统计研

究工作者不断完善统计认识方法和认识过程,在统计认识过程中,将统计认识的客观性和主观性有效地结合起来,力求使统计认识的内容与客观一致,使统计认识不断深化。

### 2.1.3 统计认识观

**系统观:**客体是由许多要素构成的一个完整系统,客体系统既有自己的运动规律,同时通过输入和输出与其环境进行着交流、保持着联系、发生着相互作用。客体系统在一定时空条件下的状态取决于该系统的结构、自身运动规律和环境影响。因此,客体系统的输入、输出、状态和结构等方面的信息客观地记录了客体的属性、行为及运行规律,统计认识主体通过统计信息加工处理可以再现客体系统,达到从整体上认识的目的。

**信息观:**信息与客体同在,物质、能力和信息是客体的三个组成要素。信息是客体的间接存在。客体的信息系统与物理系统同构(异态),统计认识主体通过对客体信息系统的研究,来认识和把握客体的物理系统。

**规律观:**客体本身在自己的运动和变化中,在自己与其它客体的相互关系中遵循着一定的规律性;客体在不同时空条件下的具体数量表现以及构成客体的各个组元的特征的数量方面大小各异,但不是杂乱无章、无规律可循的,经过大量观测研究随机性将依一定的概率趋于稳定,个别差异相互抵消呈现出主体趋势性。这种稳定性、趋势性就是统计规律性。

**可知观:**客体是可以认识的,统计认识是一个不断逼近客体本身的过程,每一个认识结果都是一定条件下的产物,都受一定因素的限制,因而统计认识是近似的、不精确的。

**逻辑观:**统计推断的前提与结论不满足逻辑必然性,表现为用前提证明结论正确的程度的一定可能性。如果说培根—穆勒的归纳法(求同法、差异法、求同差异公同法、剩余法、共变法)在颇大程度上必须以拉普拉斯的机械决定论(单值决定论)的概念

为其本体论的前提,那么在物理学中这个概念的覆灭,迫使方法论者认真注意统计方法的认识论本质。

偶然(可能消除)观:偶然性是客观存在的,既有内部组元的随机性,又有环境的干扰作用,但是这种偶然性通过大量观测可以降低、消弱。使必然性明显地突现出来。当然不可能彻底、干净地消除。

历史观:客体变化和发展具有一定的“惯性”。根据客体的“惯性”,统计认识可以从客体的过去认识现在,根据现在推知未来。虽然处在现时状态,时间的不可逆性决定了人们无法重演历史,但是信息客观、精确地记录了客体的历史状态和运动轨迹,信息系统与物理系统的同构异态,决定了统计认识可以借助客体的统计信息系统再现其变化发展过程,使历史分析与逻辑推演成为现实可能。

发展观:统计认识认为任何客体都是变化发展的,这种发展既有共时态的规模扩展,又有历时态的演进。但不论是发展还是演进,发展的动因除了内部组织的作用和环境变化的影响之外,客体的发展还有一定的惯性。且客体的发展有一定的规律性,统计认识主体可以从上述三个方面测定客体的发展方向和速度。

开放观:任何客体系统都不能独立存在,都是在与一定的环境相互作用的形式中发展的。不仅如此,客体系统必须与环境不断进行物质、能量、信息交换。封闭的系统是一个死结构,必将随着熵增而衰弱、直至消亡。系统必须是开放的、通过不断引入、吸收环境的负熵流,实现熵减,形成新的序结构,使客体系统不断注入新的活力。开放才有物质、能量、信息的交换,主体才有了认识和把握其特性的可能。统计认识就是统计认识主体与客体的信息交换。所以统计认识不仅十分注重客体本身的结构、矛盾、规律,同时十分关注系统与环境的关系和作用方式,研究环境的变化可能给客体带来的影响及其影响程度和方向,也



就是熵的输入和作用。

#### 2.1.4 统计认识过程

统计认识过程必须经历四个阶段:即获取客体信息,升华到理性认识,形成决策方案,信息反馈。

##### 一、获取客体信息

获取客体信息是统计认识主体根据统计认识宿体的目的和需求,选择科学合理的必须获取的客体信息的种类、数量以及方式方法等,在此基础上通过统计观测、实验(试验)搜集客体的信息。概括地说,就是统计主体通过感觉器官接受外部信息,对客体“可见世界”的认识。即能动反映过程的开始。这是一次统计认识的开端和前提,实现了对客体状态的描述,建构起客体的原始信息系统。从认识论来看,这一阶段相当于感性认识阶段,得到的是关于客体的状态、表象方面的认识。但是它不是纯粹的与思维无关的感知阶段,而是具有高度思维的、有目的、有意图的感知活动,是在思维的支配、控制下进行的。同时,即使是这样的感知活动有时也往往被省去了。统计认识所依据的原始信息可以不直接来源于具体实践活动的感知,而是使用他人或前人的成果,即所谓的次级资料。这种次级资料是统计认识研究的一个极为重要的原始信息来源。尽管这些次级资料归根结蒂,也是来源于人的感性实践活动,但是特定统计认识主体不需要进行实践而取得自己所需的感性材料。此外,就统计认识的方法而言,也不需统计主体重新进行实践取得感性经验,而是可以继承他人的经验,也可以通过科学思维,沿着自己发展的内在逻辑前进取得新方法。我们称之为感性认识阶段是从该阶段获得的信息是关于客体表面的、某个侧面的、零散的、局部的特征这个意义上来说的,而不过分强调信息的来源通道,即一定是统计认识主体感官对客体的感知这一规定。当然,有许多大量的统计认识研究活动,其原始信息必须经过统计认识主体亲自深

入到对象中去,获得第一手信息。

## 二、升华理性认识

升华理性认识就是根据所获取的信息,经过统计信息加工处理,获得关于客体规律性的认识。这是统计信息的第一次升华,是全新的统计信息产品,实现对客体的解释功能,达到观念上的把握。它是统计认识的重要阶段。这一阶段还可以具体化为知性认识和理性认识两个环节。

1. 统计整理是统计认识的知性认识阶段。这是理性认识的初级阶段。这一阶段的主要任务是将感性阶段获得的零碎的反映个别的原始信息进行去伪存真、去粗取精,并按一定的规则进行简单加工、分解、归类、综合、汇总,得到反映客体状态、特征的新信息。这一阶段统计认识主体将凭借科学认识能力开始对客体“可知世界”进行认识研究,使个别零碎的缺乏联系的感性对象具有秩序;但是就一般意义而言,知性可视为判断能力,知性作出的判断,只能是“非此即彼”的,其中不能包含矛盾,因为它只能认识有限的和有条件的经验事物。为进入理性认识的更高阶段打下坚实的基础。统计认识引入了概率知识,彻底克服了绝对的“非此及彼”的观念。

在这一阶段的统计认识具有三个显著的特点:

一是抽象同一性。统计整理是将统计主体采集到的个别、零碎的原始信息进行加工,使之条理化、系统化,得到反映总体特征的指标。即从一般感性材料获得抽象一般。这是知性认识的总特征。知性是撇开事物的总联系而单独地、孤立地考察事物某一方面,某一规定。黑格尔称之为“分析的理智”。它以抽象的同一性作为自己的思维原则。所谓抽象的同一性是排除差异的对立的自身同一,即  $a = a$ , 每一事物和它自身同一。黑格尔说:知性思维要求的是“坚持着固定的规定性和规定性之间彼此的差别”(黑格尔:《小逻辑》商务印书馆 1981 年,第 172 页)。一方面它坚持思维的固定性,坚持统计指标之间的严格区别,决不

容许有丝毫的空泛和不确定之处,反映的对象是什么就是什么,不能同时又是别的东西;另一方面,它排除指标、分类的灵活性、流动性,尤其是在分类、分组的时候,它只承认“非此即彼”的一面,否认“亦此亦彼”的另一面。

二是无矛盾思维。在统计整理阶段,统计主体最主要的思维方式就是无矛盾思维,它严格遵守形式逻辑的同一律、排中律、无矛盾律的原则界限,脱离了内容和发展来考察事物,追求形式上的无矛盾性。

三是可证实性。知性是“只能产生有限规定,并且只能在有限规定中活动的思维”。这些有限的规定不是空泛的,不可捉摸的,而是具有严格确定性、固定性,具有明显直观性、可靠性。统计整理所用的统计指标的内涵、计算口径、计算方法及计量单位等都是明确规定的,计算公式严格遵守逻辑运算规则。因此,用于计算统计指标的原始数据和统计整理所形成的统计指标之间保持着必然的逻辑联系,具有严密形式上的一致性,因而是可证实的。

统计知性思维具体实现三种认识功能:

一是分解混沌模糊的表象,形成抽象的规定。采集来的信息是客体的形式混沌、模糊、零碎的表象。感性认识尚不是知识,要获得知识性的信息,就必须对感性材料进行加工提炼,形成指标。知性是一种分解的能力,经过知性的分析可得到“固定的和静止的规定”,这些规定都反映客体的某一个方面。没有这些抽象规定,认识就不可能深入到客体之内部及其联系层面来把握其整体。

二是概括一类事物或关系的共同性。统计知性思维摆脱客体之具体形态,揭示一类事物之普遍共同性,而形成有限的规定。

三是把思维形式化、数量化、公理化。知性思维是严格按照公式、模式、形式进行思维活动的。它以变量分布及均值、方差、

相关等为基本概念和命题,遵守统计规则和方法进行推论,形成自己的演绎系统。

分解和计算是知性的重要认识功能。诚然,算法本身不仅是知性思维的结果,而且也是理性思维的结果,但在算法基础上的行为则是知性的。例如,计算过程分解为多项单个步骤,各个步骤、各个阶段之间具有机械的顺序性,严格的制约性,指令是以符号语言组合的形式为出发点。可见没有知性思维就不能有思维的形式化、符号化,就不能有符号形式的逻辑运算,也就没有现代的统计软件系统。

2. 统计推断和预测是统计认识的辩证理性认识阶段。辩证思维方式寓于理性的思维方式,并以扬弃的形式把后者包括在自身之中。辩证思维以知性思维为基础,在知性所建立的有限规定的基础上,寻求具有对立统一关系的具体的普遍性。辩证思维从概念和判断的动态的相互联系,往返运动中揭示它们的内容。因此,在辩证逻辑中,对于像演绎和归纳,分析和综合这样的思维方法,不满足于像形式逻辑那样把这些方法作孤立的、平列的使用,而是运用矛盾的思维方式使对立的两极相互过度、相互转化,在它们的互补统一中揭示事物运动的本质。此外,辩证思维的最主要的原则和方法是从抽象上升到具体,逻辑的与历史的相统一。具有间接性、概括性、本质性,是关于事物规律性的认识。

### 三、形成决策控制方案

形成决策控制方案就是统计认识主体根据宿体的认识目的和客体的发展变化规律,形成实现宿体目的的最优控制方案。这个阶段是统计认识的关键阶段,是统计信息的又一次升华,使抽象的理性认识具体化、应用化、形象化为指导行动的蓝图、方案、计划,达到观念地改造和创造客体,这是统计认识的价值理性认识阶段。所谓价值理性认识,就是统计认识主体借助于思维的创造力,根据辩证理性认识所发现的矛盾和提出的问题,确

定目标、形成解决矛盾的决策方案,从实际条件出发,经过试误比较,选择最优的方案,以便直接地、有效地指导变革现实。可见,统计决策认识是实现认识的第二次飞跃的准备,或者说在观念上实现第二次飞跃。因此,统计决策方案是客体的规律、环境的影响、主体的认识能力以及宿体的愿望相结合的产物。

1. 价值理性与辩证理性的区别。二者的本质区别应该是:是否结合主观愿望,不是研究“是什么”、“状态如何”,而是研究“怎么做”、“如何实现”。统计决策认识借助创造想象力把抽象思维和形象思维有机地统一起来。创造想象力是探索求知的能力,是人们根据一定目标构思现实中不存在的新模型、新图像的思维活动。在统计决策形成过程中,要先后经历二个过程,即从个别到一般和从一般到个别。前者用统计方法和统计思维将信息加工整理,得出客体规律性的认识,后者则把一般规律具体化、应用化、形象化为行动方案。

2. 统计决策认识与统计推断认识的区别。统计推断认识也要根据一定的标准选择推断认识结果,如依据标准确定参数估计值、确定最适模型、客体的未知状态或未来状态等。但是统计推断认识是在“是什么”与“不是什么”之间进行选择,是就认识的飞跃程度和深刻程度进行评价,不涉及主体和宿体的利益关系。而统计决策认识,是在“做什么”“不做什么”,“怎么做”“不怎么做”之间进行选择。不仅要看认识是否客观、准确地反映客体,而且还要考察其是否在可能的条件下最理想地解决矛盾达到最优目标,效益最好。统计决策认识的选择是价值的选择,评价原则是价值准则。所以,我们认为统计决策认识属于价值理性认识。

3. 统计决策认识是统计认识过程的一个重要阶段。统计决策认识是实现由认识到实践第二次飞跃的前提、桥梁。是完成统计认识任务,实现统计认识目的的重要阶段。从认识论意义上来说,统计认识的任务不仅在于揭示客体的状态、结构、特征、

发展变化规律和推测未知、未来状态,而且更重要的在于怎样控制、改造世界。只有在改造世界的实践中,认识的价值才能实现。如果认识结果不用于指导实践,而是空谈一阵,那么,即使统计认识对客体揭示得再深刻、解释得再明确,也没有任何意义,其价值就无法实现。统计认识指导实践需要一个中介,这个中介就是将认识结果具体化、应用化、形象化的统计决策方案。毛泽东说:“做必须先有人根据客观事实引出思想道理、意见、提出计划、方针、政策、战略、战术,方能做得好。”(《毛泽东选集》人民出版社,1969年,第1卷,第445页)。这个提出计划、方针、政策、战略、战术的任务,就是价值理性认识的内容。就统计认识来说,就是统计决策认识。通过统计决策方案,把宿体的主观目标与客观规律、现实条件紧密结合起来。使揭示事物本质特征和内在规律的统计推断认识克服间接性、抽象性在应用方面的不足,从而具有直接性、具体性、实用性的特点,对一些随机的不确定的因素采用统计方法加以量化,对客体发展变化作出概率分析,以求得对客体的变化及未来状态有更精确的统计。

#### 四、信息反馈

信息反馈是指在客体的运动过程中,将有关检验信息返回到认识主体的过程。统计信息反馈包括两个方面:一是统计信息的内反馈,具体表现为统计认识各个环节的各种统计审查、统计检验、决策方案优选以及模型模拟等。二是统计信息的外反馈,就是形成的决策控制方案在具体指导客体的运动过程中,将信息再反馈回来,进行偏误分析,并制定修正方案,使统计认识逐步逼近客观真值。统计信息反馈具体地体现了统计认识是一个不断再认识的过程。

#### 五、统计认识过程

统计认识是一个充满矛盾、螺旋式上升的无限发展过程。统计认识决不是照镜子或照像,一次完成是不可能的。统计认识的过程性主要表现在:首先,客体是一个不断运动、变化、发展

的过程。恩格斯指出:黑格尔巨大的功绩在于他第一次“把整个自然的、历史的和精神的世界描写为一个过程,即把它描写为处在不断运动、变化、转化和发展中,并企图揭示这种运动和发展的内在联系。”(《马克思恩格斯选集》第3卷,第362页,人民出版社,1995年6月)。统计认识的任务就是探索客体的结构和发展变化规律。每一次统计认识都是根据特定的信息对客体特定时空条件下的认识,当时空条件发生变化,或获得了新信息以后,就需要对客体重新进行统计研究或修正统计认识结果。因此,客体的发展变化,需要统计不断地进行认识研究。其次,从统计认识的内容上来说,它认识的是整个客体的结构和内在规律性。而整体的结构和内在规律性是许多要素综合作用的结果,它寄寓客体的种种具体现象、不同时代之中,往往呈现为主客观交织、现象与本质交错、正确与错误交错、确定与随机混合、模糊与精确包容等复杂形态。统计认识要经历一个从个别到一般,从现象到本质,从偶然到必然的运动过程。再次,从认识的本质来说,认识是主体对客体的能动反映。是一个由不知到知,由知之不多到知之甚多,由不全面到全面,由不深刻到深刻,主客体信息不断交换的无限接近的过程。统计认识也不例外。因此,作为统计认识过程来说,随着信息的不断反馈、交换,可以客观、及时、准确地认识客体,因而是至上的;但是作为任何一次认识活动,任何一次认识成果,它是特定条件的产物。属于极为有限的时空,因而又是不至上的。第四、从统计认识主体来看,作为整个统计认识主体,世界连结、世代相传是至上的,而某一统计认识主体,某一时代的统计认识主体,又是不至上的。第五、同样统计认识手段也既是至上的,又是不至上的。这种至上与不至上的矛盾,只有在世代相继的无止境的统计认识过程中才能得到解决。第六、一般与个别的矛盾运动规定了统计认识是从个别到一般又从一般到个别反复转化的再认识的过程。列宁指出:“任何一般都是个别的(一部分或一方面),任何一般都是

大致地包括一切个别事物。任何个别都不能完全地包括在一般之中。同时,(另一方面)个别和一般相联系而存在的,一般只能在个别中存在,只能通过个别而存在”(列宁:《哲学笔记》第409页)。一般只是“同实在事物的无限多的方面中的一面相符合”(《列宁全集》第27卷,第34页)。它只是认识具体事物的一个阶段,一个方面,虽然包含着绝对真理的因素,但没有穷尽整个认识对象。第七、统计检验不是一次可以完成的。统计认识不仅要进行实证性检验,还要进行实践检验。实证检验和实践检验都不是一次完成的,而是需要经历一个反复检验的过程,这是由实践检验标准的不确定性决定的。不确定性首先表现为它固有的普通性、直接现实性的优点是受历史条件限制的;其次,同作为认识主体的目的和认识手段等主观因素也有密切关系,在统计检验过程中,检验水平是由认识主体决定的,因而,同一数据,同一认识,同一方法,由于检验水平不同,可能会出现不同的结论;再次,统计检验依据的检验数据具有不确定性、随机性、个别性,用来进行检验的样本数据是偶然的,受特定时空条件的限制,除具有客体的一般之外,更具有显明的个别性。实践标准具有相对性、不确定性,因而不能无条件地完全地证实或驳倒任何理论或观念,我们决不能把实践标准绝对化,以为只要经过一次检验就能完全证实或驳倒某种理论、观念,不需要反复检验。同样,我们也决不能把实践检验过的原有的知识、认识绝对化,把它当作为不需要重新认识、不断完善、不断发展的永恒真理。

不仅从整体上说,统计是一个认识过程,就是在一次统计认识活动中,也存在着统计再认识。如迭代优选求解就是一个不断再认识的过程,是认识向客体运动的辩证性。

统计再认识是新旧认识联系的中介,即新旧认识的肯定与否定的辩证统一,是认识的连续性与非连续性的统一,是人们深化认识、发展认识的必要环节。通过再认识这一环节,把认识从片面性、不深刻的本质向全面性、更深刻的本质不断推进。



统计再认识是对原有认识的否定,但不是全盘的否定,不是任意的否定,肯定原有认识中正确的、合理的东西,而否定其不正确的或不完全正确的东西,从而深化认识、发展认识,意味着再认识是一个总结经验、消化经验的过程。

再认识体现“认识向客体的运动”的辩证性:它表明认识不是一次完成的,而是要经过一个不断反复,由低级向高级发展的无限过程;它表明认识过程不是一帆风顺的、平静的,而是充满着新与旧、真理与谬误、主体与客体等一系列矛盾斗争。它体现了认识主体的能动性,即人们在实践中检验真理、修正错误、完善和发展认识的能动作用。

#### 六、统计“测不准现象”

在统计认识过程中,往往出现“测不准现象”。应该说不论是统计认识主体,还是统计认识客体都希望统计认识客观、准确、精确。但是,这只是主观愿望,客观上是很难做到的,甚至可以说是不能做到的。在系统的构成要素中,任何一个微小组元都具有不确定性,这是固有的一种属性。从一个系统来看,它不仅受内部各个组元的不确定性的影响,而且还受到环境因素的不确定性的影响。此外,在统计认识过程中还存在统计认识手段与统计认识客体的“中介”现象。当统计认识主体利用某种统计手段,在自己的认识图式中对客体的特征进行整合时,依据的不是完全的、纯粹的客体要素的信息,而是统计信息采集手段与客体要素之间相互作用表现出来的现象。统计主体无法准确地判断哪些是来自客体本身的、真实的信息,哪些是受统计认识手段影响而产生的干扰性的噪声信息,这时客体属性也表现为不确定性,出现“测不准现象”。可见,“测不准现象”从表面上看是统计结果与统计认识客体的不一致,完全是统计认识主体认识能力造成的。事实上,“测不准现象”也是客观的,是统计客体存在和发展的内外因素、统计认识手段的社会历史性等客观条件决定的,不是主观自生的,它是对主客体相互作用形成的一种

关系状况的反映,这种客观关系和状况不依赖于统计认识主体的主观意识而转移。当然,“虚假统计”另当别论。

我们说统计认识具有不确定性,“测不准现象”是客观的,但是,这种不确定性并没有否定组元的客观性。主体因素渗透到认识中去,但其中一部分的主体因素是可以客观化的,例如,主体的思维方式、知识结构、测量仪器等等是智力因素的物化形式,是人类认识成果的凝结物,是可以使之客观化而加以测定的。

### 2.1.5 统计认识的客观性和相对性

统计认识是一个多种信息的综合体,包括客体场信息、客体系统环境的干扰信息、主体内神经结构中先已凝结着的信息、主体神经系统之外的体内其它方面的干扰信息等等。因此,统计认识的客观性决定于客体信息场的中介建构作用和认识与客体特性的某种对应性;主观能动性来自于生理结构、认识结构以及认识手段对认识的中介建构作用。

统计认识作为一个过程是对客体无限逼近的过程,是无条件的。但就每一次认识活动来说是有限的、有条件的。因为主体总是在一定环境、一定认知结构中去认识客体的。“一定环境”给客体加上了一个“约束”,这是一个外部约束,使客体的多样性受到限制;“一定的认知结构”是主体的“内部约束”,主体的认知结构给客体的多样性又增加了一个约束。此外,还有认识手段的约束。

统计认识的主体相对性。客观世界具有普遍相互作用的客观本质性,既规定着事物的存在方式,也规定着认识主体的认识方式。人是客观世界的一部分,只能在客观世界之中,只能在遵循客观规律的基础之上,并且只能以人自己的方式来认识世界。主体以自身状态(包括认知结构、选择方法、物化工具)本身规定着自身认识的限度。即主体是通过自身这个参照系来对客体进

行把握的,参照系对每个人来说不尽一样,同一主体不同时间、条件下也有变化。因此,对同一物的认识不会完全一致。不仅感性认识如此,由于理性认识活动只能是在主体认知结构所提供的思维模式的规范下进行,因而也具有相对性。另一方面,主体的相对性还表现在认识相对于人的历史的、社会的认识水平和程度。

统计认识成果的相对性。客体是可以认识的,这是绝对的;任何一个具体的认识都是不完全的,这是相对的。从辩证唯物主义来看,统计认识的真理性取决于关于客观真理的学说,即取决于相对真理与绝对真理的学说。统计认识所反映的内容(即客体的某一属性)在现实的认识中是变化着的,这个被反映的内容对客观实在的符合是一个过程。这个过程的辩证法在于映象(术语的含义,陈述的内容、意义)同客体符合的矛盾性质。一方面这个客体(绝对真理)无条件地反映在映象中,反映着对世界上的情况的真实理解,这一点是无条件的。但另一方面,这个反映是不完全的、近似的,在这个意义上是相对的(相对真理)。由此得出结论,在每一个有关客体的统计认识中都带有绝对真理(映象同客体之绝对的、完全的符合)的“内核”、成分,和不确定的部分即相对真理(映象客体之近似的、不完全的符合)。统计认识的相对性事实上就是人类知识的可变性和发展性的具体化。科学知识在历史上每一个确定的时期除了用确定性和绝对性表征以外,还要用映象同原型的近似程度的某种不确定性和相对性来表征。统计认识与客观事实的关系就象两条渐近线,一是真值,二是统计值。两条线一齐向前延伸,彼此不断接近,但是永远不会相交。两者的差别就是统计误差。统计误差使得统计认识并不无条件地直接就是统计认识客体,而现实也不直接就是统计认识结果。因此,统计认识结果(指标体系或模型)一方面与现实不是等价的,而另一方面按其现实的关系来说,也不是虚构的,而是在某种程度上、从某个方面把现实包括在自

身的内容之中。

### 2.1.6 统计认识系统

统计认识是一个完整的系统。这个系统是由统计认识主体、统计认识客体、统计认识中介和统计认识宿体四个基本要素构成的。四者相互影响、相互作用,缺一不可。而每个要素各自又形成了一个子系统,具有各自的结构和功能。

#### 一、统计认识系统是一个自控系统

从控制论的观点来看,统计认识系统也是一个自控系统。控制有多重意义:从人与工具的关系看,控制为掌舵术、驾船术、管理艺术。从人与机器、能量的关系上看,控制就是机械化的操作。从人与信息的关系来看,控制就意味着在人的活动和机器的活动中如何获取、处理和运用信息。从普通控制论来看,控制指的是自控系统通过信息选择、传递、反馈,把行为和动作施加于被控制对象,使系统的特征和变化保持在目的规定的限度内,以达到正常而有效地发挥功能,从而把被控客体引入某一预定的合乎目的的状态,即从概率性较大的状态引向概率性较小的状态。可见,实践活动也具有信息意义。因为实践不是盲目的,总是在标志着主体目的性的理论信息指导下进行的,这个主体信息控制着实践方向、途径、速度、广度、深度,在这个意义上说,实践也是主体信息对客体积极作用的过程。正是物质性活动的实践具有信息意义,信息方可作为认识控制论系统的一个决定因素。

控制论的控制有两个本质特征:一是控制就是对信息的控制,二是控制过程具有能动性。控制系统就是信息自控系统,控制过程就是信息加工处理和使用过程。

通过对信息的控制来认识和改造外部世界,是当今人类所能达到的最有效的控制方式。苏联学者索波列夫指出“控制原理的实质在于,巨大质量的运动和行动,巨大能量的传递和转

变,可通过带有信息的不大的质量和不大的能量来指挥和控制”(《控制论的若干基本特征》载《国外自然科学哲学资料逻辑》第四辑,上海人民出版社,1965年版,第93页)由于信息具有不同于质量和能量的特点,如它可以识别、转换、存贮、传递、再生、共享等,所以,我们可以运用信息的传递来取代巨大质量的运行、巨大能量的传递和交换,从而大大提高人类对外部世界的控制能力和控制效果,以便更精确地认识和把握客体发展的规律性。

在控制过程中,人们根据获取的信息,编制出适宜的指令,通过反馈调节,保持系统的外部环境的平衡状态,同时人们通过掌握干扰信息或系统的状态信息,并把它转换成抵消或纠正偏差的控制指令,排除干扰,以保证系统目的的实现。在现代,离开了信息,人类就寸步难行。更谈不上控制外部世界。

此外,由于客体内在矛盾性以及外部因素的干扰,规定着它的不确定性。这种不确定性表现在事物的发展存在多种可能,多个方向,即它处于多种可能性空间,正是事物处于多种可能性空间之中,才有必要选择。要实现控制,认识主体对多种可能性空间进行分析,然后根据自己确定的目的,在多种可能性空间中选择其状态,或某一状态目标,以便把可能性空间缩小到一定的范围,从概率性较大的状态引向概率性较小的状态。向可能性空间的缩小,则标志着认识主体能动性的发挥和控制能力的提高,标志着人们的认识更加接近客观真理,通过选择,缩小可能性空间,排除干扰,消除混乱,明确方向,使系统从不确定到比较确定状态,这是控制目的性的表现。

从控制论的观点看,统计认识主体与统计认识客体相互作用的实质在于对信息的控制。也就是说统计认识过程是主体对客体之间的双向信息控制的过程。同时,也是主体的自我控制过程。主客体的双向控制指的是一方面统计认识主体按照统计认识宿体的目的和需要等具体要求对客体信息的获取、储存、传递、加工、处理、提供、反馈等思维操作活动,以便观念地把客体

移入头脑,这是统计认识主体运用已有的认识图式、认知结构去同化客体信息的过程;另一方面客体也要控制主体,制约主体。当认识主体不能用先存认知结构、认识图式同化客体信息的条件下,主体要顺应客体,改变或调整原有的认识图式、认知结构。同时,主体还必须不断进行自我控制,包括对认识过程非理性因素诸如价值取向、道德观念、行为规范以及情感意志、心情等心理状态的自我控制和自我调节。

## 二、统计认识具有认识控制论的基本特点

1. 统计认识是从可能性空间出发的。这是认识控制论的一个基本特征。可能性空间的观点克服牛顿经典力学的直线性、机械性。牛顿经典力学把客体描述为质点,没有内在矛盾,只有必然性,没有偶然性,一切都是纯粹的;在运动形式上,物体只能作直线式的运动,只有一种可能,一个方向,一切都是确定的;在发展变化上,未来一切都是由过去的一切决定的。这种观点在认识论中的表现就是形而上学地理解认识过程,这正是旧唯物论认识论机械性、直线性的自然科学根源。从可能性空间出发,强调发挥主体能动性,能动性又受可能性空间制约。统计认识强调客体由于内在的矛盾和环境的干扰,客体具有不确定性;这种不确定性又不是杂乱无章的,而是具有一定的规律性的。客体的不确定性决定了客体的发展变化存在着多种可能性。这些各种可能性就构成了客体的可能性空间。统计认识就是在客体的可能性空间内排除偶然性、不确定性,探求客体的必然性和规律性。统计认识不仅在可能性空间把握客体的特征、状态和发展趋势,而且得出的认识结果也是概率认识而不是“非此即彼”的僵化认识。有时为了研究客体的某些特征,在获取信息时要进行统计试验(实验),在进行统计试验时,通过控制试验条件,将客体的可能性空间缩小到一定范围,以便在繁杂的信息中获取所需要的信息。

2. 统计认识具有明确的目的性。在统计认识过程中,统计

认识主体处于一定的意识准备状态或认识定势,它是认识主体先存各种意识状态的综合统一体。从结构上说,就是头脑中已具有的认知结构,具体包括知识结构和评价结构,它是过去认识活动的积淀物,是在实践→认识的反复过程中而建构起来的,并经过不断的检验、修正而逐步完善和发展。它是认识主体控制客体信息的重要工具,是主体选择和确定认识对象的框架,是主体对客体信息进行过滤和整合的规范和准则。贯穿于整个统计认识活动全过程的主体对客体信息的选择功能是统计认识目的性的又一具体表现。从统计认识主体在可能性空间确定搜集哪些信息一直到制定统计控制决策方案,无时无刻不在进行选择,选择正是认识目的性的体现。此外,认识的目的性还体现在统计思维具有自我调节自我组织的功能。

3. 统计认识注重研究客体的动态过程。统计认识并不否认静态研究(即在共时态下认识研究客体的状态结构)的重要性,但是,统计认识现在的目的在于把握客体的未来行为。所以,研究客体的运动轨迹,揭示其变化规律,推断未来状态以及使客体的发展状态符合宿体的目的要求,即将客体的发展变化引导到宿体需要的特殊状态的统计控制决策方案等统计认识方法都是以研究客体的动态过程为基础的。

4. 统计认识系统的信息反馈机制。在控制论中,信息反馈是一个极为重要的概念。控制过程中,信息反馈又是主控者对受控者施行控制的重要手段和环节。统计认识系统具有明显的信息反馈机制。参数估计、构建客体的各种统计模型等统计认识方法,都是在客体的所有可能性空间中,不断通过信息的反馈,逐步缩小可能性空间,使统计认识主体可以在更小的可能性空间范围内对客体作出概率性认识。没有不断的信息反馈就无法步入优良估计,不可能得到适应模型,研制不出来最优统计控制决策方案。统计认识有效地利用了信息内外反馈机制,从而使其具有显明的不断再学习的特点,成为人类认识的普适性工

具,认识结果具有了无限逼近客观真实的可能性。

统计认识系统的内外信息反馈是相互结合的统一体。并且这种信息反馈贯穿于始终。从统计认识过程来看,统计认识主体将不断根据客体发展变化的信息修正完善统计认识结果,每一次再认识都是信息反馈引起的,没有得到新的、有价值的反馈信息,再认识就失去了意义。这是统计认识系统的第一层信息反馈。在每一次认识过程中,统计认识主体,都要根据信息加工处理中的新信息及时完善起初设计的统计认识活动方案,这可以说是统计认识系统的第二层信息反馈。就一种统计认识方法来说也具有再学习功能,通过信息反馈机制,使认识结果逐渐逼近客观真实。如在采集信息阶段,要相继进行多种方式对所获得的信息进行确认;移动平均,指数平滑等是最典型的再学习方法;参数估计要根据每一次估计的误差信息不断搜索;适时预报和统计控制决策方案的优化等都是一个信息反馈过程;这可以说是统计认识系统的第三层信息反馈。此外,需要特别强调的是,在统计宿体根据统计认识结果以及其它依据在组织具体实践活动中,即所谓的“事中”,统计认识及时的信息反馈适时校正施控行为,使得实践活动不断缩小目标差,稳定地实现预设的目标。

在宿体对客体的控制过程中,统计认识主体把握客体及其周围环境的变化,运用统计方法排除(准确地说是减少、降低)各种干扰的影响,提高统计认识的有序度和内稳态,使统计认识朝着预定的目的方向发展。在具体信息操作中,又要根据统计主体的价值取向和传统文化框架,根据宿体的认识目的等要求,借助认知结构对信息进行过滤和综合、变换以及内外反馈,从而实现主体对客体信息的控制,观念地改造和创造客体。从可能性空间出发,明确认识的目的性,研究事物的行为方式,正确地进行信息反馈等基本观点,则是主体对客体信息控制的内在机制,它体现了认识论的时代精神和特征,并以此而别于以往的认识



论。统计认识论把认识能动性具体化、深刻化,揭示认识能动性的根本观点就在于发挥认识主体性,实现主体对客体未来状态、发展态势、调控方案的认识和控制,从而规定统计认识论必须面向未来,着力测定客体的未来状态和发展趋势。从本质上说,统计认识论是精确地测量、计算客体特征、结构和规律的认识论,是科学地预见、控制未来的认识论,超前的认识论,它从更深层次、更具体地发挥认识主体的能动性,是马克思主义认识论的具体化。

## 2.2 统计认识主体

### 2.2.1 统计认识主体的质的规定性

统计认识实践的主体,不等同于生产实践的主体。统计实践的主体是在社会体力劳动和脑力劳动分工的基础上出现的科学劳动者的一个组成部分。统计认识主体的职能是通过大量观测或实验、试验,取得客体的原始信息,实现对客体的初步信息建构,并经过统计思维的综合,掌握客观现象的因果联系,找出客体的运动变化规律。因此,统计认识主体是探索新的、未知的客观规律实践的主体。从认识论来看,统计认识主体是指有头脑、能思维并且具有一定统计知识结构和能力的从事统计认识活动的个人和群体。

所谓“有头脑、能思维”是指作为统计认识主体的人,不是任何形式抽象的精神实体,既不是费希特的“自我”,也不是黑格尔的“绝对精神或绝对观念”,而是有血有肉有生命的物质实体,马克思在《1844年经济学—哲学书稿》(《马克思恩格斯全集》第42卷,第167页,人民出版社,1965年第1版)中明确指出,“主体是人”,而人是“自然的、肉体的、感性的、对象的存在物”。这就是说统计认识主体是作为自然界一部分的人。人只有作为生命的

自然存在物这个自然前提,才能成为统计认识主体,成为统计认识活动的物质承担者。可见,这是对统计认识主体的最基本的规定,不具备这一特征不可能成为统计认识主体,而具备了这一特征也才具备了成为统计认识主体的可能,是一个可能的主体。第二个规定就是作为统计认识主体的人必须“具有一定统计知识结构和能力”。这是统计认识主体的特殊规定性。统计认识是人类认识的一个方面,一种途径。因此,统计认识具有特殊性,特殊的认识活动形成了专门的认识经验、认识理论、认识方法以及认识手段。这些专门知识同化并纳入主体原有的认知结构,表现出特殊的认识能力。同时具备这两个规定的主体,仍然是一个可能的主体。要成为现实主体,必须同时满足第三个规定“从事统计认识活动”。这一规定首先强调,统计认识主体是实际从事统计认识活动的现实的社会人。马克思早在《1844年经济学——哲学手稿》(《马克思恩格斯全集》第42卷,第89—103页,人民出版社,1965年第1版)中就指出,人的本质就是人的感性活动。后来,马克思和恩格斯又在他们合著的《德意志意识形态》一书中明确指出:“我们不是从人们所说的、所想象的、所设想的东西出发,也不是从存在于口头上所说的,思考出来的,想象出来的,设想出来的人出发,去理解真正的人。我们的出发点是从事实际活动的人。”(《马克思恩格斯全集》第3卷第30页人民出版社,1965年第1版)。从事实际活动的人必然要受到其所处的历史条件、自然环境、社会关系、意识形态、自然观、发展观、认识观、科技水平等等的制约。因而统计认识主体是一个社会主体。同时,还强调具备统计认知结构和能力的社会人,只有在其实际从事统计认识活动中,才成为统计认识主体。因此,统计认识主体又是一个现实主体。这是对统计认识主体的质的规定。

除上述三条规定之外,统计认识主体还有一个显著的特点,就是统计认识目的的社会性。一般来说,认识主体进行认识的

目的就是为了使自己对客体有一个更深入的了解和客观的把握,以便有效地控制对象系统的运行,更好地满足人类发展的需要。但是,随着统计科学的发展和人类认识的深入,统计认识已逐步发展成为一个专门的领域,分离出来一批专门的群体,形成了一个专门的产业。因而统计认识主体从事统计认识的目的往往不是为了满足自己的需要,而是为了满足另一个认识主体或更大的认识群体的需求。

一般而言,主体与客体关系包含三个层面:即改造与被改造,认识与被认识,索取与被索取。相应地也就有实践主体性,认识主体性和价值主体性。统计认识主体与统计认识客体主要是认识与被认识的关系。因此,对统计认识主体来说,具有显著的认识主体性。

### 2.2.2 统计认识主体的形式

统计认识主体的形式是多种多样的,从其活动的结构来看,主要有个人主体,群体主体和社会主体。

个人主体就是指从事统计认识活动的主体是个人,个人在一定的社会环境和认识条件下,相对独立地进行统计认识活动。显然,这种认识活动的认识能力、认识结果等具有明显的个别性,不可避免地会受到其所处的客观环境和主观条件的制约。同一认识客体对于不同的认识个体来说,其认识过程、认识深度、以及形成的概念系统等会有明显的不同。尽管如此,但是个人主体的存在是必然的、普遍的。这首先是由统计认识的普遍性决定的。统计认识作为以信息为依据以测度为手段的一种重要认识方法,普遍适应于认识自然、社会、人类思维等各个领域和科学研究、人类活动的组织管理等各个方面。不论是在生活中、还是在工作中,遇到的每一个对象、每一个问题都需要了解本质,把握特征和规律,制定行动方案和措施,而在认识活动中,几乎都要自觉不自觉地、或多或少地运用到统计认识方法,统计

认识思路、统计认识评价标准、统计认识结果等。特别是随着科学技术的发展和信息化进程的加快,统计认识将成为每一个行为者获取信息、把握问题、控制系统的重要手段之一,也是一个行为者认识能力和认识水平的一个重要标志。其次,个人主体的存在是群体主体和社会主体存在的前提,离开了个人主体,群体主体和社会主体就成了一个空洞的概念,失去了客观、现实的形式和内容。此外,随着信息网络的普及,信息越来越规范,越来越完备,越来越丰富,信息处理技术越来越先进,大量的复杂的信息处理将由计算机自动完成,信息共享越来越广泛和深入,个人主体将有广阔的生存空间和发展前景,将发挥必不可少的作用。

群体主体是指从事统计认识活动的主体是一个群体、一个集团,具有一定的组织形式和组织机构,共同协作从事统计认识活动,实现认识功能。如统计局、信息公司、咨询中心、调查公司、统计研究所、统计事务所等。群体主体是若干个人主体的集合,因而在一定程度上摆脱了个人的局限性。通过优化结构,形成大大高于个人主体质级的强大的系统功能。这是统计认识主体的主要形式。

社会主体从理论上讲是指主体是一定的社会。但在统计认识活动中,社会主体是指有共同认识目的,并以同一认识客体为研究对象,从事统计认识的所有个人主体和群体主体的综合、总和。它比群体主体更具有综合性;拥有全社会的力量,集合了整个社会从事统计认识活动者的智慧;凝结了全社会的认识能力;是所属时期的最高认识水平。因此,其整体优势是其它主体形式无法与此相比的,是认识主体的最高组织形式。

个人、群体和社会三者是紧密联系相互制约的统一体。这一方面表现为个人和群体都是一定社会的细胞和组成部分,都只有在社会中并通过社会,才成为主体的。另一方面社会主体也不能脱离具体的个人,否则就成了一种抽象物。此外,还表现

在个人认识、群体认识、社会认识三者之间的相互影响,相互渗透,相互制约。具体说来就是社会认识、群体认识个体化和个人认识群体化、社会化的双向运动。社会认识、群体认识个体化是指社会主体和群体主体的认识成果的内容及其认识规范影响个人主体,并取得个人主体的认同,进而被个人主体接受、吸纳、整合于自己的认知体系,形成自己的思维定式的过程。通过这一过程,使统计认识成果全面推广,使全社会的每一个人主体的认识水平跃上一个新层次,提高人类认识世界的水平。个人认识群体化、社会化,是指个人主体的认识成果的内容及其思维定式影响群体主体和社会主体,并取得赞同、承认和接受的过程。这是个人对社会的贡献。

如果从统计主体认识活动的目的来看,统计主体分为两种情况:一种是既是认识者,也是使用者。即既是统计信息产品的生产者,也是统计信息产品的使用者,表现为主体与宿体统一。另一种仅仅是认识者不是使用者,而是专门根据用户需要从事统计认识活动,为用户提供统计信息产品。后者是随着统计认识活动的发展而分离出来的,是统计认识方法技术的专门化、复杂化和统计认识范围的不断扩大以及统计认识结果使用的社会化的必然产物,并将越来越占主要地位。

### 2.2.3 统计认识主体的结构

不论是社会主体,群体主体还是个人主体都是一定要素构成的有机体,都是一个系统。个人主体和群体主体都是社会认识主体系统的子系统;个人主体同时也是群体认识主体的子系统。并且个人主体也是一个系统,它是由个人的认识器官(包括其延伸物),个人的认知结构,个人的客观环境和社会条件等要素构成的一个统一体。各个要素相互作用,相互制约,共同决定着个人认识主体对客体的认识程度和认识成果的质量。

统计认识主体作为一个系统,从其实现认识功能的角度来

看,是由激发驱动子系统、整合导向子系统和调控子系统组成的。激发驱动子系统是统计认识主体自觉意识的有机结构,具体包括对象意识和自我意识两种意识要素。统计认识主体的对象意识是指主体能够自觉地在自己意识中把外部世界中的特定事物作为自己活动的对象,形成关于外部对象的对象意识,从而把主体与客体区别开来,并对认识对象给予明确的界定。统计主体的自我意识是由主体的自我认识意志方面的自我调节和自我控制,情感方面的自我体验等因素构成的。是对统计认识主体本身的生活和活动(物质和思维的活动)的意识,主体对自身的需要、属性、状态和活动及其同外部客体的关系的意识,就是自我意识。主体具有的对象意识和自我意识是主体形成自觉的需要意识前提。主体具有需要意识并能时时对它进行反思,从而产生驱动激发力量。

激发驱动子系统是统计主体进行统计认识的内在驱动力。它决定统计主体要对何种事物进行认识?何时认识?认识的内容及其广度和深度等。显然,这种内在驱动力就是需要意识。换句话说,当统计主体意识到存在某种需要时,便产生了对某种目标的追求,以便使需要得到满足。需要意识愈强烈,统计主体的认识活动就愈具内在的驱动力。这里需要特别值得注意的是,对需要的正确理解。虽然需要是主体对客观事物的需求在头脑中的反映,并在人的意识中表现为主观欲望,但需要是具有对象性的东西,是进行科学研究、组织管理等具体要求,是一种客观的规定,它不是由主体的主观意志决定的任意的东西。没有客观的对象或离开了人的实践活动的前提条件,就不会产生主体的需要。统计认识主体需要的产生归根结蒂是由改造客观世界和进行科学研究的实践决定的。实践变革客观对象,一方面要求在变革前对变革对象的性质、特征、规律及其与其它事物的内在联系等有比较深刻的认识和精确的把握,另一方面在变革过程及其结束后,使人们能了解和掌握对象之属性并通过实

践活动来规定人的需要,这就是统计认识主体需要的客观性。需要意识是人对自身需要的意识、反思。统计认识主体的需要意识具有显著的社会性。社会需要决定主体的思维方向,个人需要意识只有反映和适应社会需要,才能正确选择,确定自己的行为 and 动机。统计认识主体的社会性还表现在统计认识往往不是为了主体自身使用,而是为了其它行为主体,乃至全社会共享。需要是决定统计主体行为方向的一种内在驱动力即动机的根据,是统计主体在外界环境、主体的自身生理状态和活动的刺激下产生的。

统计主体的整合导向子系统是统计认识主体进行统计认识活动的的能力结构,主要包括主体的心理结构、主体的知识结构、范畴系统和逻辑推理规则等四个方面的内容。

主体的心理结构是主体认识能力的自然基础。是在人类种系进化过程中通过遗传方式而获得的,是自然——社会的产物,具有社会历史性。

主体的知识结构,是主体认识能力的核心。主体的知识结构,从其来源来看,又可细分为理论和经验两部分,即间接知识和直接知识两部分。间接知识是先人的直接知识,是在一定社会历史条件下和社会环境中,通过后天的学习和实践、社会交往和继承而形成的。从内容来看有关于对象系统的知识、统计知识、信息处理技术知识等。

范畴系统是统计认识的工具。列宁指出,概念、范畴系统都是作为客观世界的反映而被表现出来的,是事物的辩证法创造观念的辩证法,而不是相反。概念、范畴体系是人们在认识和改造客观世界的实践活动中所形成的,是物质世界及其关系的反映。概念、范畴一旦形成,它就获得认识的功能,成为主体认识事物的工具,人类认识史上所取得的第一个成果都是用概念、范畴的形式固定下来的。统计主体在进行统计认识的过程中,必须从业已形成的概念、范畴出发,并进行细化和量化,才可能进

行信息搜集和信息加工,即使是最后的统计认识成果,也需要借助于一系列概念进行记录和表述。

逻辑推理规则是统计认识的逻辑手段,是统计认识主体在统计认识活动中的行为规范。

上述四个部分的有机结合,形成了统计认识主体的能力结构并内化为认识图式。在统计主体的认识活动的不断重复过程中,逐渐失去了它们的感性的外衣而具有形式化、理性化,成为统计认识主体完成一种认识所必需的思维框架和心理状态。在认识过程中起着整合导向的作用,它直接地影响和制约着统计主体接收信息、加工处理信息的感知过程和思维过程,影响和制约着统计认识主体把握世界的广度和深度,并作为主体认识定式,控制着认识过程的方向。

调控子系统是统计认识主体进行统计认识的非理性因素的影响力。有了认识的驱动力和认识能力之后,统计主体的认识活动还不可避免地要受到一些非理性因素的影响和制约。如情感、意志、信息、理想、习惯、本能等。这些非理性因素直接调节、控制着统计主体的认识活动。

可见,统计主体的认识结果是统计认识主体的内在驱动力,认识能力和调节控制因素综合作用的结果,三者缺一不可。因此,要提高统计认识水平,必须在提高统计认识能力的同时,强化统计认识主体的内在驱动力并改善调节控制因素,三者合理优化,才能顺利完成统计认识活动,获得良好的认识成果。

#### 2.2.4 统计认识主体的功能

选择定向功能。就是统计认识主体,根据需求和自身结构对统计认识客体、属性之取舍以及思维空间的定向的机能,是统计认识主体得以获取客体信息并规定认识活动方向的客观机制。选择是统计认识过程的必要环节,离开了选择,认识过程就难以实现。选择贯穿于认识过程的各个环节和各个方面。首先



确定统计主体之需求,必须进行选择。在任何时候,对任何客体,都有进一步认识的需要,但是统计认识主体永远也不可能满足各种需求,必须根据各种需求的迫切程度和重要性以及自身的认知结构选择一种需要,确立一个目标。当确立了某种需要之后,就必须对认识客体进行选择 and 界定。可是客体有各种各样的特征和属性,统计认识主体又必须选择这次认识活动要认识的主要特征和属性,以及认识的切入点和要达到的目标。为了达到所确定的认识目标,又需要选择统计指标体系,以及搜集信息的方式、方法和处理信息的技术,提供信息的手段等等。从不同认识层次来看,在感性认识阶段,观测和实验都需要进行选择。复杂的统计总体具有无穷的信息,统计认识主体不需要也不可能把客体所包含的全部信息都接收过来,而必须进行筛选;然后选择适合的统计方法经过去粗取精,去伪存真,由表及里的信息加工过程得到理性认识。在价值理性认识过程中(优化决策方案过程中),需要对各种理论、目标、方案进行选择,确定一个整体最优方案。

同化整合功能。从信息论的观点看,统计认识的过程实质上就是统计认识主体借助于认识图式对信息的同化整合的过程。统计认识主体经过系统地、有选择地获得的客体原始信息,将输入到统计认识系统的加工系统。主体认识图式提供了同化客体信息的规范、方法和准则。然后按照这些规范、准则借助恰当的统计方法对原始信息进行同化整合。这个同化过程不仅是信息净化过程而且是信息变换过程,即经过对信息的加工,从整体上把握信息,系统地把它们吸收到原有的认识图式中去,又以认识图式为依据,把信息纯化,使其失去感性形态,转化为模型、概念,取得观念形态。之后,认识的行程进入概念的整合阶段。这是一系列概念、范畴,规律的建构,达到对客体本质的把握和规律性的认识。

控制调节功能。统计认识主体是一个自组织系统,具有自

我调节,自我控制的功能。首先表现为自我调节认知结构。当客体信息能被同化时,原有的认识图式只需进行局部修补、完善或微小调整就能完成认识任务,达到认识目的。当客体信息不能被同化时,在主体对客体信息进行思维操作的反馈、变换的相互作用中,主体就要发挥自我调节、自我控制的功能,修改、更新原有认知模式,实现原有认知结构向新的认知结构的变换,建构新的认知结构。这种结构的修正、变换、重建是通过认识主体自我调节功能来实现的。其次,自我调节认识目标。统计认识主体不但要确定认识目标,而且有时要调整认识目标。当目标定得过高或过低时,认识主体必须根据认识需要,认识能力,认识手段以及客体的具体情况等及时进行调整。再次,自我调节统计认识方向。统计认识是一个极为复杂的过程,要排除一切各种偶然因素的扰动,从大量随机性中得出统计规律性来。因此,统计认识也可以简单地称之为消除随机性的过程。在具体操作中,往往表现为试误过程,即经过不断尝试,使得误差最小的过程。试误过程具有相对性。在试误过程中必须经常调节方向。

## 2.3 统计认识客体

### 2.3.1 统计认识客体的属性

统计认识客体简单地说就是统计认识对象。从认识论来看,统计认识客体可以概括为:那些纳入统计认识主体活动范围,为主体对象性活动所指向的客观事物。纳入主体的认识活动范围是指:①是预先客观存在着的事物;②具有满足主体需要的属性;③在实践上按人的方式同物发生关系;④发展、强化统计认识主体自身的本质力量,即提高和加强主体认识和改造世界的能力。客观事物能否成为认识客体,一方面取决于它的客观存在性,同时还取决于认识主体的认识手段和认识能力。因

而在不同的历史发展阶段,统计认识都有自己不同的对象和范围。同时,统计认识客体是有复杂结构的、同其他事物相互联系、相互作用的动态对象。

根据上述统计认识客体的哲学定义,统计认识客体具有如下属性:

客观性。指认识客体是客观存在的,有其固有的内在客观规律性和发展变化规律,不以人的意志为转移。自然客体和社会客体的客观性不言而喻,在此仅就思维客体的客观性进行简要探析。

首先,思维成果是人脑的产物即物质的产物,思维过程是信息输入→加工整理→输出的客观过程。

其次,思维成果是对物质实践活动的反映,思维过程依据客观规则。

再次,思维客体是主体认知结构的客观产物,主体的认知结构决定了主体对客体的反映,这种认知结构是客观的,认知结构一旦建构,主体对客体的反映事实上也就产生了。因此,看似主观的东西,实际上是客观的产物。

最后,从精神客体的作用来看,它是客观的,精神产品一旦产生,它就客观地发挥着影响作用。

对象性。指统计认识客体不是自在的东西,而是客观世界中纳入统计认识活动范围的同统计认识主体活动有功能联系而被具体指向的对象。对象性是统计认识客体的本质特征,是统计认识客体之所以成为认识客体的现实条件。客观存在只是潜在的客体,只有同统计认识主体建立了对象性关系的客观事物才是统计认识的现实客体。如人口毫无疑问是客观存在着的,是可以作为统计认识对象的潜在统计认识客体。但只有进行人口普查或其它统计认识活动时,才纳入统计认识活动的范围,才从潜在的统计认识客体变为现实的客体。显然统计认识客体同统计认识主体是对立统一体,二者相互依存,互为前提。没有统

计认识主体,就根本无统计认识客体可言,同样没有统计认识客体,统计认识主体也就没有存在的实际意义。

此外,统计认识客体的对象性,从另一个侧面说明统计认识是一个开放的系统。随着科学技术的发展和人类实践活动的需求,统计认识主体的能力不断增加,新的统计认识客体将不断产生,统计认识客体的范围将不断扩大,统计认识的领域将不断拓展。

自发能动性。统计认识客体的自发能动性主要是指它具有自己运动的内在源泉和动因。它是客体运动及其规律性的根本决定因素。毛泽东同志曾经指出:“唯物辩证法认为外因是变化的条件,内因是变化的根据,外因通过内因而起作用。”(《毛泽东选集》一卷本,人民出版社,1964年4月版,第277页),因此,统计认识要了解和掌握客体的运动状态和发展变化规律,就需要对统计认识客体运动的内在动力源泉进行认识研究。在这个意义上说,对客体状态的统计认识是表面的、肤浅的、暂时的、手段性的,对客体自发能动性即动因的认识才是本质的、深刻的、永久的、目的性的。只有对客体的动因(自发能动性)有了把握,才能真实客观地认识客体运动规律性。

客体的自发能动性不仅是客体运动的动力源泉,同时还决定了客体与主体、客体与环境之间的相互作用关系。这就告诉我们,在统计认识活动中,不仅要深入研究客体本身,还必须认真研究客体与环境的内在关系,结合环境变化,掌握客体运动、运动规律及其发展变化趋势。客体与主体的相互作用关系说明统计认识不是单项式的,(即统计主体有能动的方面,没有受动的方面;或客体只有受动的方面,没有能动的方面),而是一种双向式的互动。

系统性。统计认识客体是一个完备的系统,是由许多要素组成的一个有机整体。客体系统的结构性,要求统计认识必须研究客体的组成要素及其位置和作用。同时客体系统的整体

性,要求统计认识决不能停留在结构研究、要素分析的层面,系统的整体性质和功能决不是各个要素性质和功能的简单机械总和或叠加,必须认识系统的整体性质和功能。黑格尔曾经说过“一个活的有机体的官能和肢体并不能仅视作那个有机体各个部分,因为这些肢体器官只有在它们的统一体里,它们才是机体的器官,它们对于那有机的统一体是有联系的,决非毫不相干的。只有在解剖学者手里,这些官能和肢体才是些单纯的机械部分。但在那种情况下,解剖学者所要处理的也不再是活的身体,而是尸体了。”(黑格尔《小逻辑》第282页,商务印书馆,1980年7月第2版)。即使是主要研究某个要素的功能,也必须清醒地认识到部分的功能是在整体性的前提下发挥作用的。如人的手具有操作工具的功能,但是当人的整体性能不好时(疲劳、精神不振、生病等),手的操作功能将明显减退,直至完全消失;更为甚者,当手与身体分离之后,将完全失去一切原有的功能。

相对独立性。就是说统计认识客体的各个组成要素、各个层面、各种属性、各个发展阶段有其自身的结构、状态、属性、功能和发展变化规律。就各个要素来说,相对于其更高一级的系统结构,它是一个组成要素,但就其本身来说,也是一个完整的系统,高一级的系统可以看作是它的环境。因此,在共时态,系统的各个要素、各种属性可以间隔开来,可以分解开来单独进行观测研究。在历时态上,相对于某一时间点系统具有客观的、现实的仅属于这一时间点的状态,统计认识主体可以掌握它,并通过对一系列离散的状态进行研究,认识系统的运动轨迹以及规律性。统计认识客体的相对独立性,为统计主体认识客体的结构和运动规律提供了客观依据和现实可能性。

承认认识客体的相对独立性,不是对客体的系统性的否定。二者是对立统一的。正如墨西哥科学家托马斯·阿·布罗金指出:“物理自然界尽管有其完善的统一,至少暂时地可以分解为隔离的部分。假如不是这样,我们甚至不得不在一个极简单的

认识情况下考虑到宇宙的全部要素,因而对自然过程的任何研究、任何理解都将永远是不可能的……”(转引自[苏]拉扎列夫等:《认知结构与科学革命》湖南人民出版社,1986年版,第66页)。相对独立性为人们认识客体提供可能性,而系统性则是人们综合地、深刻地认识客体的特性、结构、功能、规律的基础。

社会历史性。统计客体的社会历史性是指统计认识客体随着人类社会的变化和发展而发展。首先表现在统计认识客体是一定社会历史阶段的产物。一方面一些客体仅属于特定的历史阶段,另一方面属于人类历史长河的东西在不同的历史阶段上有其不同状态和特性。其次,统计认识客体是由统计认识活动决定的。客观事物是否成为统计认识客体取决于是否对其进行统计认识研究,只有纳入统计认识活动范围,对象化了的那些客观事物才是统计认识客体。统计认识活动是人类在一定社会历史条件下进行的,因而统计认识活动及其认识客体具有明显的社会历史性。随着人类实践水平和统计技术的提高,会有越来越多的对象被用来进行统计认识研究而成为统计认识的客体。因此,在进行统计认识研究时,必须紧密结合特定的社会历史条件,统计认识结果也是特定社会历史条件下的产物。对于这一点,社会客体、思维客体以及自然客体中的人化自然客体都是不言而喻的。天然自然客体的社会历史性乍看起来似乎有一些矛盾,但仔细分析,就不难理解了。天然自然虽然具有自然属性、自然规律,但是它之所以能够成为统计认识客体,是因为它的某些属性已经取得了社会意义,也就是说与人类社会活动紧密结合为一个有机整体,不可分割,在相互影响、相互作用下各自发生变化和发展。而这个社会意义是人类实践赋予的。人类实践活动的社会历史性决定了自然客体具有明显的社会历史性。

可测量性。就是说统计认识客体只能是原则上可以观测到的。这里有三层含义:一是被认识客体可以直接或借助仪器间接作用于观测者,二是某种观测及其相应的工具和仪器在技术

上可以实现,三是其特征可以量化。统计认识的显著特点是从量的方面研究对象的性质、结构及其运动规律性,因此统计认识的可测量性是其能够成为统计认识对象的决定性条件。对于那些在目前科学技术水平,测量手段等条件下,尚无法进行量化或无法进行观测的对象,不论其如何重要,人类如何急需,也不能用统计方法进行认识研究,因而也就不能成为统计认识客体了。这里需要指出的是,可测量性是相对的,不是绝对的。一方面客体的可测量性是相对于特定时期的科学技术水平和测量手段以及人类在实践活动中对其的控制能力和控制程度、方式来说的。换句话说,是受制于人的认识水平和控制能力。就本质而言,应该说任何客体的可测量性是绝对的,不存在不可测量的客体。因为客体是客观存在的,不是主观臆造的,并且具有质和量两个方面以及运动的根本属性,人们可以在客体的运动中对其运动轨迹进行测量和记录,从而进行统计认识研究。另一方面可测量性有测量程度、测量精度、测量方式之别,有的可以测量得非常精细、准确;有的则仅可以实现量化分类,因而所运用的统计认识方法和统计认识结果的深度有所不同。对于只可用定类尺度测量的客体来说,我们只能通过大量观测,获得结构比例、众数等统计信息。随着科学技术的发展,人类认识器官将无限延伸,测量手段越来越科学,将有更多的对象具有可测量性而成为统计认识客体。

可认识性。客体的运动变化既有确定的一面,又有不确定的另一面。从其内部矛盾所规定的发展方向、途径来看,客体的行为是确定的,有一定规律可循。但从客体的可能性空间来看,又是不确定的,存在各种可能的未来状态。从统计认识来看,统计认识承认任何客体的运动都是有一定规律性的,这种规律性是由客体内部矛盾性决定的,具体表现为客体系统运动的惯性,或者叫做客体系统的“动态记忆性”。与此同时,强调客体系统的现在状态不完全决定于历史状态的输入,在运动过程中不可避免

要受到一些随机因素扰动的影响。这些随机因素有的来自客体内部各组成要素,有的来自客体所处环境条件的变化。因此,统计认识主体对客体动态过程的认识也只能是概率性的把握。

### 2.3.2 统计认识客体的形式

就现阶段来看,统计认识客体根据其性质不同,有自然客体、社会客体和思维客体。

自然客体就是纳入统计主体对象性活动范围的自然界的客观事物。具体包括天然自然和人工自然。前者是指天然存在的自然物,即自然界本身固有的、未经人类加工的天然物质资料。后者是指人化自然即人类在生产活动中形成的自然界以及人的物质机体。人化自然从其物质属性来看,是天然物质资料的转化,是改变了形态和属性的天然物质;从其劳动属性(价值)来看,是人类劳动的凝结物是人类智力和体力的结晶。随着人类认识世界和改造世界的能力的加强,人化自然将愈来愈普遍、广泛。这些人类对象性活动的产物又作为他人,或另一次认识活动的对象。至于人的物质机体,勿用多言,它是人的自然属性的一面。正如恩格斯说的,我们躯体上的肉、血和头脑都是属于自然界的一部分。显然属于自然客体的范畴。

社会客体是指纳入人的实践活动范围并与统计认识主体发生联系的社会关系即人类生产实践活动及其结果——人们的社会关系和政治关系。目的是为了认识和利用经济状态及其发展变化规律和社会规律(包括政治、法律关系等)。

思维客体是指思维活动或思维操作的产物以及对象化的思维活动本身。思维活动的成果表现为精神产品,是脑力劳动和体力劳动分化的产物,是具有相对独立性、有其自身发展规律的主观存在,是客观存在的主观映象。从形式上看带有主观特征,但内容是客观的(是具有主观形式的客观)。在科技和经济高速发展的现时代,精神形态的知识、信息,以及人的思维活动过程



已经成为极为重要的统计认识客体了。

### 2.3.3 统计认识客体的结构

任何一个统计认识客体都具有一定的结构。所谓结构具有两层含义,一是指统计认识客体内部各组成要素之间的相互联系、相互作用的方式即各要素在空间的排列和组合的具体形式和具体关系,这可以称之为静态结构。二是指统计认识客体在不同时间的各个状态之间存在的相互影响、相互作用的方式,称之为动态结构。

静态结构是对统计认识客体的共时态考察。静态结构决定客体的本质、属性及其存在形式,是一事物区别于它事物的决定因素(根据)。客体的静态结构首先表现为客体的组成要素,也就是说客体是由多少要素组成的、是由哪些要素组成的。其次表现为各要素所占的份额。由种类相同的元素构成的客体由于其内部各自的份额不同就会有不同的性质。如一氧化碳和二氧化碳,都是由两种元素构成的,这两种元素都是碳和氧,但是由于其份额不同性质截然相反。再如混凝土都是沙子、石灰和水组成的,但是其比例不同,性质就会有明显的差异。社会现象也是如此,比如各种所有制所占的比例不同,决定社会制度有所区别,社会主义社会公有制为主体占绝大比重。当公有制这个主体地位失去以后,社会主义性质也将不复存在。再次表现为各要素之间的关系,即相互联系、相互作用的形式和程度。组元固然重要,是最基本的构件,没有组元或组元残缺、质劣,当然不会有良好的结构,但是,关系也极为重要,这些要素是依靠某种关系联结在一起的。没有关系,这些要素将是些独立的分散的东西,各个组元的质量再好,也都是一些“游兵散将”,形不成整体或不含有良好的属性。就象有了优质的钢筋、水泥、砖、沙石等建材,没有按一定的关系将它们联结起来就形不成完美的建筑物,永远也是一些建材。一个事物或称一个系统,只有同时具有

要素和关系这两大构件,才有了结构,这两大构件的优劣决定了客体的优劣和性质。从客体的整体性、系统性来考察,关系比要素更重要,关系决定着各个要素的功能的发挥,这种决定性既可能是促进,也可能是制约。从而决定了客体的整体功能可能大于各要素功能之总和,也可能等于,甚至小于各要素功能之和。

此外,统计认识客体的静态结构还表现在各组成要素在空间上的排列和组合形式。由于统计认识客体往往是一个复杂的大系统,因而构成客体的各个要素的空间布局也比较重要,如研究人口问题时,人口的分布情况,研究商业网时,网点的分布情况等都是重要的内容。更进一步考察,统计认识客体的结构还表现为层次性。即被认识客体是由许多要素组成的,其中的任何一个要素又是由许多组元组成的,依次下去,从形式上看,客体的结构呈现为树型网状结构,层次清晰分明。因此,在对客体进行静态考察时,既要弄清其基本组成要素及其在空间上的排列和所占的份额,更要掌握其相互之间的关系,只有这样,才能把握客体的全貌。

动态结构是对统计认识客体的历时态考察。任何事物都是运动的、发展的。运动是有规律性的,这种规律性就表现为客体在不同时刻上各个状态之间的相互影响和相互作用。有些客体的动态结构比较简单,客体在某一时刻的状态与其历史(以及)的行为无关,体现为无记性即没有惯性,每个时刻的行为是随机的,这样的客体其发展变化随机性占主导地位。有些现象则不然,它在某时刻上的状态受制于两个因素,一是其过去的状态,二是现时刻受到的扰动,即表现为对过去的记忆。记忆有长短、强弱之分,因此,呈现出各种各样的动态结构。研究认识客体的结构时,精确测定和把握客体的动态结构是非常重要的。根据其动态结构和运动规律性,结合历史状态以及在未来时刻可能发生的某些重大影响,人们可以对客体的未来状态、走势作出预测,也可通过调控某些影响因素,使其未来的状态到达人们的预

期目标,实现控制功能。

上述探讨的是统计认识客体的内部结构,事实上在进行统计认识研究时,还必须考察客体的外部结构。外部结构也就是客体存在的环境,若将统计认识客体视为一个系统,用系统论的语言来表述,外部结构就是系统环境。辩证唯物主义认为任何一个事物都不会独立地存在,它必然与它事物发生联系才能存在和发展。这种“与其它事物的联系”就是客体的外部结构。从客体结构的层次性来考察,在树状结构中,每一个节点,当将其纳入认识活动对象化了的时候,都可以成为一个统计认识客体。从这个节点(客体)出发,往下发展形成了一个树型结构,这个节点(客体)就是这个树型结构的根。就象树根必须在土地里通过各种须根与土地里的各种成份保持一定的联系,树才能成活成长一样,这个节点(客体)往上一层看,它又是更大系统的一个组成要素,必然与同层次的其它要素相互制约、相互作用。因此,我们可以认为该系统的环境(客体的外部结构)就是更高一层次系统的内部结构。

外部结构不仅决定了客体的存在,同时还决定了客体功能的发挥,并赋予其发展的动力。世界上任何一个客体都是在一定的环境中存在和发展的,没有独立存在的现象,这就如同植物在一定的环境中生长,生命在一定环境中存在和发展一样,离开了它的“一定环境”,就不能存在和发展,就必然死亡和消失。外部结构的内容具体表现为客体系统与其环境的物质、能量和信息交流。集中到一点就是客体系统与环境的熵流。耗散结构理论表明,如果一个系统失去与环境的熵流,客体将随着熵增而“热寂”。所谓熵是描述由大量分子组成的热力学系统客观性质的一个物理量,可以推广到一切由大量元素或子系统组成的巨大系统。若令  $S$  表示熵,  $S_0$  表示初始状态对应的熵,  $Q$  为热量,  $T$  为绝对温度,按克劳修斯的定义

$$S - S_0 = \int Q/T$$

为简化起见,令  $S_0 = 0$ , 则熵公式为  $S = S \mid Q/T$

可见  $S$  是一个相对量, 是系统微观层次混乱程度的一种度量。“是一个系统失去了的‘信息’的度量”(陈润生《熵》载《百科知识》1987 年第 10 期), 维纳在《控制论》中明确提出信息就是负熵的命题, 并把信息与系统组织性联系起来, 指出: “信息量的概念非常自然地隶属于统计力学的一个古典概念——熵。正如一个系统中的信息量是它的组织程度的度量, 一个系统的熵就是它的无组织程度的度量, 这一个正好是那一个的负数”。(维纳:《控制论》科学出版社, 1962 年版, 第 11 页) 维纳认为“信息这个名称的内容就是我们对外部世界的调节并使我们的调节为外界所了解而与外界交换来的东西”(维纳:《人有人的用处》, 商务印书馆 1978 年版, 第 9 页)。在这个基础上布里渊提出了负熵(negentropy)概念, 代替薛定谔讲的“负的熵”(negative entropy)。他把负熵原理表示为信息  $I =$  熵  $S$  的减少 = 负熵  $N$  的增加, 这说明, 获得信息要以消耗负熵源的负熵为代价。

热力学第二定律的数学表述是熵增加原理: 当物质系统经绝热过程, 由一个状态到达另一个状态时, 它的熵不会减少, 熵在可逆绝热过程中保持不变, 在不可逆绝热过程中增加。用  $ds$  表示熵的增加, 则熵增原理可表示为  $ds \geq 0$ 。也就是说, 在一个没有与环境物质、能量、信息交换的系统(孤立系统)中, 对于各种可能的变动来说, 平衡态的熵最大。当这种系统处于非平衡态时, 有  $ds > 0$ , 系统朝着熵增加方向演化, 达到平衡态时,  $ds = 0$ , 系统保持在平衡态, 最大熵态就是稳定平衡态。换句话说, 对于同环境没有物质能量交换的系统, 不能形成“活”的有序结构, 必然走向最无序的状态。只有在与环境广泛交换物质、能量, 引入负熵流, 且引入的负熵在绝对值上大于客体内部产生的熵增时, 客体总熵逐步减少, 使客体由无序趋向新的有序, 形成一种“活”的高度稳定有序的耗散结构。这是将来统计认识的一个重要方面, 是统计认识深度的一个质的飞跃。

此外,现象与本质,形式与内容的关系也是统计认识客体结构的重要内容。

#### 2.3.4 统计认识客体的功能

这里研究统计认识客体的功能,不是考虑每一个客体在人类活动中所具有的特殊功能,而是考察其作为统计认识客体,在统计认识活动中所具有的功能。也就是它作为构成统计认识系统的一个子系统,在整个统计认识系统中所具有的功能,这些功能是所有统计认识客体的共性。

我们强调统计认识是认识主体与认识客体相互作用、相互影响的一个过程,在这个过程中统计认识客体处于被认识、被研究的地位,但是它在统计认识过程中不完全是被动的、受动的,也具有一定的内在的、客观的能动性。

客体的激发认识需求的功能。统计认识活动是有需求趋动而产生的,是有目的指向的。这个认识需求是以客体的客观存在为前提,由客体的性质、作用决定的。可以设想一个根本不存在的客体,人们也就根本不会有要认识它的意识。同样,一个当前对人类无足轻重的客体,人们也不会产生对其进行统计认识的强烈欲望。没有统计认识需求,也就无统计认识可言。

概略地说,是客体的客观存在及其在人类活动中的地位和作用产生了统计认识需求,统计认识需求决定了统计认识主体的认识行为。统计认识的目的是了解和掌握客体的性质、结构和运动规律。可见客体不仅具有激发认识需求的功能,还有决定统计认识目的的功能。

客体的信息源功能。统计认识活动是统计信息生产活动。统计认识目的是通过统计认识主体运用统计方法对客体的信息进行采集、加工、开发、传递、提供实现的。作为一个信息活动,首先要有信息,而要有信息就必须有信息源。显然统计认识的对象是统计认识客体,因此统计活动的内容——信息当然是来

自统计认识客体,也就是说统计认识是统计认识主体和统计认识客体之间的信息交换。没有或没有足够的、客观的来自客体的信息,就不能对客体有客观、深刻、全面的认识。

在统计认识活动中,统计认识主体关注的是统计信息源,即如何从信源获得反映客体本质特征的信息。随着统计信息加工方法和技术的发展,加工、开发信息不是统计认识的关键问题。决定统计认识成果优劣的关键是来自信息源的数量和质量。所以说,在统计认识活动中,统计认识主体关注的不是统计认识客体本身,而是客体作为信息源的角色。因为统计认识既不是为了实际占有客体,又不是实际操作客体,而是为了在观念上把握客体,是通过信息操作,经过主体与客体进行信息交换来实现认识客体的目的。

客体的决定认识方法的功能。表面上看,用什么样的统计方法来研究客体是统计认识主体选择的。事实上统计认识主体是根据认识客体的性质、结构及其发出的信息进行选择的,也就是说统计认识主体所做的选择并不是主观的、随意的,而是受制于统计认识客体,是由统计认识客体决定的。不同的客体有不同的运动形式,有不同的性质,有不同的规律,统计主体只有选择相适应的统计方法,才能获得客观认识结果。例如,为了获得一个地区的人均寿命或新生婴儿的性别的信息,人们只能用观测的方法,而不能用试验(实验)的方法;对于获得的定类信息,只能计算比例,只能用众数反映一般水平,而不能计算算术平均数和中位数等等。

## 2.4 统计认识手段

### 2.4.1 统计认识手段的基本属性

统计认识活动是统计认识主体与统计认识客体之间的信息

交换。只有统计认识主体和统计认识客体,仍然还无法实现二者之间的相互交换。二者之间的信息交换必须通过一个中介才能实现。统计认识的中介是沟通统计认识主体和客体、实现二者相互作用的中间环节,是统计认识主体用以反映客体的条件和手段的总和。在认识过程中,认识主体和客体的相互作用是被多级中介着的,客体本身无论如何也不能进入主体、能够进入主体意识的仅仅是已在诸多中介中几经变换、选择、建构过了的客体的信息。

统计认识的中介系统可以描述为:客体→客体信息场→(主体神经系统的)生理结构→(主体内在凝结的)认知结构→认识工具(物化手段)→认识方法(统计操作、运算和推理的规则)→主体。统计认识客体和主体就象一条河流的两岸,必须通过桥梁才能联结起来,才能实现彼此间的交流。我们把统计认识活动中,置于统计认识主体和统计认识客体之间,实现二者之间信息交换,起联结、沟通、交换作用的一切中介的总和称之为统计认识手段。它包括客体信息场、主体的生理结构、认知结构、统计认识方法、统计认识工具、信息符号、信息载体以及统计认识方法和认识工具的使用方式、信息输入输出传递方式、信道、社会文化背景等等。没有起中介作用的统计认识手段系统,统计认识过程、统计认识目的就无法实现,统计认识需求就无法满足。

#### 2.4.2 客体信息场

统计认识的第一个中介环节是统计认识客体信息场。信息场就是指客体辐射或反射出来的载荷信息的粒子场。物体由于内部和外部的客观的相互作用,不断地向外辐射或反射粒子,这样就将自身的存在方式和状态的信息由这个具有特定的性能和分布的粒子场传送了出来。这就揭示了产生信息的根源存在于物质世界本身之中,存在于物质的相互作用之中,存在于物质自

身运动之中。正是物质的这种自身显示的属性,才使这个世界成为可知的世界。视觉的中介是物体辐射或反射出来的光子场,嗅、味觉的中介是辐射出来的各类分子场,听觉的中介是客体振动出来的机械振动波场,触觉则是以各类热温场、机械力场或化学递质场为中介的,这些不同的场只是客体某些方面的信息载体,而绝不是客体本身。我们能看见太阳的形色,但不是太阳进入了我们的视觉系统;我们可以闻到遥远地方的花香,也不是花进入了我们的嗅觉系统。实质上在主体认识客体时,主客体之间没有直接接触,而那些直接接触的刺激物却并不能成为这一过程中的客体,它只能扮演向主体传递另一物的信息的载体角色。换句话说,我们永远只能借助于第三者来把握我们的对象。不仅如此,直接接触感官的信息载体物也是无法进入客体的。这里涉及到信息传递和转换的一个基本特征:同一信息可以在不同的载体物中传递。刺激物和感官之间完成的是某种能量的转换,根据生理研究,认为外界一切性质的刺激物都转换成神经元的电脉冲以及神经元间相互联系的突触间的化学递质的活动。通过这种转换,刺激物将自身载负的客体信息传递给各类感受器,在这里,真正进入人之感知系统的是客体信息,而不是什么具体的物。

与此同时,主体也在不断地向外辐射或反射信息场。在主客体相互作用的中介(场)上,同时存在着互逆的两种信息流的运动。一种是主体信息向客体方向的运动,一种是客体信息向主体方向的运动。这两种互逆的信息的运动使客体和主体都会发生某种相应的变化。主体正是在客体信息作用引起的自身变化(生理结构,认知结构的变化)中完成对客体把握的认识过程的。在这一认识过程中,主体不但对认识对象有所把握,而且同时也必然对自身的某些特性(如认识能力状况等)有所认识。

就统计认识客体信息场的产生过程而言,它本身就是在客体内部或客体与环境物之间的相互作用中建构出来的。客体信



息场是通过场结构内部的差异性以及这个场结构与环境物结构的差异性显示客体本身的内部差异性,以及新客体与环境物的差异性的模式的。虽然,场的内外差异性和客体本身的内外差异性保持着某种同构对应的模式不变性,但是,这两个差异性的具体形式却是并不完全同一的。这种具体形式上的差异是由于客体借助了一个与自身性质不同的物来显示自身而造成的。如,物体反射的光子场并不等同于物体自身原素的性质,物体振动引起的在介质中传播的机械波并不等同于物体自身振动的性质,而在不同的介质(空气、液体、固体)中传播着的振动波的形式仍然是有区别的。如此,客体信息场中的信息模式固然是显示客体的存在方式和状态的,但是,这种模式显示的形式却又无法摆脱载体物性质的规定。信息模式显示的形式并不是从客体显示出来的,而只能是在中介场的传输中建构出来的。在这里,我们有理由把客体信息的具体呈现样态看成是客体的某些特征信息方面和载体物的某些特征信息方面的综合。如能更详细地考察的话,在这一信息综合的过程中,势必还存在着环境中其它方面的信息场信息的干扰渗透。事情就是这样,任何物的信息在其产生出来的同时,就已经综合着另外一些物的信息方面了。

### 2.4.3 主体生理结构

第二个中间环节是主体自身的神经系统(神经细胞、感受器、各种神经通路、脑皮质)的生理结构。在生理上高度特异化了的人体神经系统就是我们的认识赖以产生的神经生理结构。生理学研究表明,神经生理结构的差异是生成一般精神活动上的差异的最为本质的基础。每个人的神经系统都与所有其他人的神经系统在生理结构上具有某种具体的差异。由于存在着这种具体的差异,所以我们每一个人在对同一对象的同一方面的信息进行感知和处理时的具体感受在实际上是并不完全相同的,色盲与正常视觉的区别就是一个极端的例子,其实,所有的

人对同一种颜色的具体感受程度都是有区别的,只不过由于这种视觉差异的区别的恒定性,以及人们只有通过对应符号的约定才能交流所看见的信息,所以,这种视觉上的差异在人与人之间又是无法进行比较的。但是人们比较容易区别出由于神经生理结构的不同所造成的感觉阈值上的差异,如近视者,耳背者、灵鼻子等。

通过神经系统的生理结构,认识主体对所接受的外界信息进行了某种适应自身神经生理结构特征的建构。这种建构无论在信息映现的模式上,还是在这种模式的呈现方式上都普遍存在。由于感受器的特异化,神经结构对外界信息的接收首先是经过了某种特化选择,但是,这些被选择接收了的外界信息一旦进入神经系统,便会被转换成适应这个神经结构特征的形式。这些形式并不是从外界信息中选择来的,而是在外界信息的刺激下由神经结构的特性建构出来的。如视觉系统视见时,人们往往只是把握了对象图形中的某些有差异的点,通过对这些点信息的把握建立某种扫描路径,然后再在先已建构起来的经验图式的产品背景中建构出关于对象图形的整体模式。就这一过程来看,建构出来的对象图形的整体模式往往要比真正进入视觉的外部信息多一些什么,或少一些什么。有趣的是,不同的人对同一对象的视觉其实是十分不同的,因为他们根本不可能完全重合地感觉到同样的一组差异点,更不可能通过完全同一条扫描路径把这些点组构起来,更何况每个人先已建构起来的经验认知图式又总是存在有某些差别。另外,对对象图形把握的视觉总是依赖于视网膜神经细胞的不停顿的旋转,神经细胞的这种活动方式的任何差异都将给认识的图形带来某种细微的差别。

由于神经生理结构的差异,同样的外界信息刺激便会产生出完全不同的具体感受体验。这种感受体验的不同只能用中介建构的规定性来解释。

神经系统的生理结构对信息的建构还受神经系统之外的体内其它系统信息的干扰和渗透。因为人的神经系统的活动是在与体内其它生理系统的相互作用中展开的,这种相互作用同样会把某些神经系统之外的体内信息输入到神经系统的活动之中。

神经生理结构对认识的作用也是一个动态过程。一方面从神经生理结构变化来说,它既与人体的发育变化有关,又与人的认识经验有关。另一方面从神经生理结构的活动来说,它有其自身的活动方式。人的认识必须借助神经生理结构本身的某种活动才能展开。如听觉要靠耳膜的振动才能完成。任何一种外界信息的接受、传递、加工、改造都必须通过神经细胞本身的电传导和神经细胞间的化学传导的活动来完成。在这个意义上说,神经生理结构是一个具有自身特定活动方式的、不断变化、发展着的动态结构系统。

#### 2.4.4 主体认知结构

第三个中介环节是主体先已建构起来的认知结构。在人的认知结构中,凝结着的并不仅仅是属于知性范围的认识图式、框架等,在这个知性的认知结构之上,还凝结着属于理性认知结构范围的思维模式、方法。人所接受和认知的这个外界信息在这个理性认知结构的中介下被重新改造,建构,从而创造出一个超越具体对象世界的人类理性的、文化的世界。人类理性认识对感性、知性和具体对象世界的这种超越的性质,充分展示着人类认识的信息综合建构、信息主体创造的深层本质。在人的认知结构中还渗透着其它一些心理因素,如,情感、意志、价值、审美、潜意识的活动等等,这些因素同样是认识产生的中介因素,同样参与了认识产生的信息建构过程。

必须提及的一点是,由于理性认识能力的发展,使人类的感性、知性活动也起到了升华,在人们认识过程中,先已建构出来的认知结构的诸因素(感性、知性、理性;情感、意志;价值、审美;

潜意识和意识等)是互相渗透、融为一体的。并且,人的理性的认知结构又对其它较低层次的认知结构施以某种控制。由于人的认知结构的这种整体统一性,使认识产生的中介建构过程呈现着极为复杂多面的、多重信息交互作用、渗透综合的特性。

从发生学的角度来看,这个主体先已建构起来的认知结构是以往感性经验的信息、理性创造的信息、继承他人的信息的凝结。这个认知结构先是经过感知学习和思维的认识过程获得,然后又作为一个已经完成的主体内部条件而先已构成了。被接受的客体信息正是在这个认知结构的参照背景的中介下被主体明晰感知或加工处理的。个体认知结构上的差异,以及处于不同发展阶段上的同一个体的认知结构发展状态上的差异,都将使这个感知的具体图式、加工处理的具体过程有所不同。认知结构更是一个不断进化、完善的过程,随着学习和实践的增强和深入而不断变化。认知结构的变化是绝对的,而不是相对的。从某种意义上说,每一次认识都是新的认知结构中中介下进行的。就统计认识而言,作为中介的认知结构包括两个层次的含义:一是非统计认识主体即一般认识主体的认知结构,二是统计认识主体的认知结构。在这种意义上说,统计认识本身就是一个再认识即是在一般认识主体的认识基础上,对客体的深入把握。当认识客体为自然现象时,统计研究往往是在相关专业人员利用具体专业知识和手段获得初始认识的基础上进行的;当认识客体为人和社会现象时,统计认识是在被研究的人或被调查的人的个人认识的基础上进行的。所有这些专业人员、被认识的人以及被调查者的认知结构形成了统计认知结构的第一层含义。第二层含义即统计主体的认知结构,包括统计理论、思维方法、有关客体的实质性理论和统计认识经验等等。

#### 2.4.5 统计认识工具

第四个中介环节是主体认识的物化手段。统计认识中的工

具不同于一般的生产工具。统计认识实践的工具有的特点,表现在统计主体为了达到某种测量、描述的特殊目的,采用不同的研究方法和各种科学仪器设备。目的和方法的有机统一,体现了统计实践工具的认识意义。现代科学仪器已经从感官的延长发展为思维器官的延长。统计认识主体对客体的认识、建构,必须借助于专门的科学仪器和工具。一方面通过仪器和工具,扩展和强化了主体感官上的机能,使主体能够在更大范围和程度上去观测、感知认识对象;另一方面便于进行精确地研究。

要认识特定的客体,要认识特定客体的特定层次、特定方面的特性,就必须应用与之相适应的认识工具、方法和途径。这就是说,特定的客体、特定的属性只有在特定的主体面前才能呈现它的客观性、它自身存在的真谛。这个特定的主体就是具有健康的神经生理结构和良好的认知结构并掌握了相应的认识手段的人。特定的主体只能在某些具体样态上,只能从一定深度上把握某些客体的某些层次上的特征,而其它一些事物,以及这些事物的其它一些层次上的特征对这一主体来说是隐蔽的。这就是同一对象,甚至是同一原始信息,不同认识主体的认识存在着差异的根源。仁者见仁,智者见智就是这个道理。

统计认识的物化手段乃是主体生理结构和认知结构的延伸和外化,是直接客观化了的主体认识条件,是主体的参照系的一个特定部分。物化手段正如人之感官,它只能对客体的某种特性进行相对探测。复杂仪器是比人之自然感官更加精密化了的人之感官。它只能比人之自然感官更为特异化,可以将人的认识引入人的自然感官不可能达到的客观领域,使人的认识逐步向客观世界的真谛趋近。处于人之肉体感官可直接感受的客观世界之外的种种信息,只有通过工具中介的选择、变换、改造和建构,才能使我们认识。

统计认识工具是统计认识手段的硬件,是物质性的。它是依据统计认识主体的需要,按照自然物的属性和规律对自然物

进行加工改造而成的工具。它的作用是强化主体的认识器官,扩大主体的认识范围,增强主体的认识能力。心理学反射理论将肉体器官分为主体效应器官、感受和思维器官。与此相适应,作为人类器官延伸物的统计认识工具也可分为三类:一是延伸主体效应器官的统计认识工具,主要是在统计实验或试验中使用的部分仪器、设备,如干涉仪、拌料器、模拟机等,它们在认识过程中可代替人的肉体器官作用于客体,通过改变客体的外部形态或内在结构,简化、纯化客体,使客体处于极端的或特殊的状态,模拟和再现客体等方式。二是延伸主体感受器官的统计认识工具,主要是指监测、测量、记录仪器设备,如测量噪声的纳声器,测量农作物等用的遥感器,以及显微镜,温度计等等。它们的重要功能是代替统计认识主体对统计认识客体的信息进行采集,也可称为统计信息采集工具。这类工具克服了统计认识主体本身感受器官的生理局限性,使那些不能为统计主体感官直接感知的事物和属性可以感知。从客体的表层深入到感知客体的结构和规律,从大致地感知发展为精确计量从而扩大了统计认识对象的范围,提高了统计认识主体的感觉能力和接收信息的能力。三是延伸主体思维器官的统计认识工具。主要是指进行信息运算、加工、推断的机器,如电子计算机,智能机(机器人)等,它们在统计认识活动中代替人脑进行思维操作,智力存贮(方法系统)、信息存贮、加工、逻辑推理等,因此,也可称为统计信息加工、存贮工具。

#### 2.4.6 统计认识方法

统计认识方法就是统计认识中的操作、运算和推理规则。统计方法渗透于整个统计认识过程。诸如认识工具和仪器的操作与使用、语言符号所指代的语义、以及信息和加工整理等环节都必须按照一定的规则和程序运作,否则,统计认识将无法进行。实际上,统计方法作为统计主体的思维工具是一种没有被

物化的思维工具。统计认识方法和统计认识工具都是统计认识主体用来使主观诉诸于客观的手段。如果没有它们的“接引”或“引渡”，认识的目的就将永远停留在主观不实的此岸，而达不到客观现实的彼岸。如果说二者的区别，那么认识工具是以实物形态存在的探测工具，而认识方法则是以观念形态存在的认识工具。

统计认识工具加上巧妙的统计认识方法，统计主体就能产生出探索客观规律的创造力量。

统计认识方法是统计认识手段的软件，是精神性的。它是人类思维延伸而形成的。具体包括统计信息采集方法、整理方法、传递方法、加工方法、存贮方法、提供方法等，因此，也可称之为思维工具。统计认识方法以一定的物质（如纸介质、磁介质）为载体而存在和交流。强调统计认识方法的精神性是指统计认识方法是统计认识主体的精神产品，是统计认识主体思维的延伸，虽然以一定物质为载体而存在，但是人们在认识过程中对它使用的着眼点并不在于其载体、物质外壳，而在于它的精神内容，因此，称之为统计认识手段的软件。

不论是统计认识手段的硬件还是软件都具有客观性、社会历史性、中介性、共享性等属性。

客观性。这是统计认识手段的基本属性。统计认识工具和统计认识方法都是在统计认识主体之外而客观存在的，不依赖于统计认识主体的意志而存在、变化和消失。对于统计认识工具来说，其构成材料是物质的，其存在形式是物质的，因此其客观性是显而易见的。对于统计认识方法来说，它借助一定的物质载体而存在，这是其形式的客观性，作为精神性的统计认识方法不仅其客观来源，而且一旦形成就有了不依赖于统计认识活动的具体主体的意志为转移而独立存在，成为人类共同的社会财富。从其功能和作用形式看也是客观的，不依赖于统计认识主体的意志而转移，被加工的统计信息一经确定，运用某种方法

对其进行加工,不管主体的意志如何,希求得到什么结果,也不管是由谁来加工,其结果是客观的。如对给定一组数据,求其集中趋势,离中趋势等统计特征,无论谁来计算都是一致的(计算错误是另一回事)、客观的。

社会历史性。不论是统计认识工具,还是统计认识方法都不是从来就有的,都是随着人类活动和科学技术的发展而产生并分离出来的,都是一定社会历史条件下的产物。是受一定历史条件制约的,是人类认识成果的凝结物。都具有鲜明的时代特征。

中介性。作为联结统计认识主体和统计认识客体的统计认识手段,同时具有主体和客体的属性。从统计认识主体方面来看,统计认识手段是统计认识主体认识器官的延伸,是统计认识主体创造的,是统计认识主体劳动成果的结晶,凝结着主体的目的因素并在发挥作用过程中实现着主体的目的和意志,属于主体的一部分。从统计认识客体方面来看,统计认识手段是统计认识客体信息向主体传递的通道。在这个意义上说,统计认识手段也是统计认识客体的放大,因而统计认识手段必须具有统计认识客体的特征,与客体保持同构,符合客体本质属性。但是,统计认识手段既不是统计认识主体,也不是统计认识客体,而是主体与客体之间具有二重性特征的中介。

共享性。统计认识手段,特别是统计认识方法可供所有统计认识主体乃至全人类共同享有和利用。统计认识方法是全人类统计认识知识的结晶,从时间的角度考察是历代人不断扬弃的结果;从空间研究,是国内外相互借鉴、综合的结果,属全人类共同拥有的知识。作为方法本身来说,可连续使用或同时运用于不同的各个客体。统计认识方法将一代一代传下去,不会因某个统计认识主体使用而消耗掉或损耗,引起使用价值磨损和价值降低。绝大部分统计认识工具,同样也具有共享性。如采集信息的勘测卫星,处理信息的计算机、人工智能机和传递信息



的信息网等等。

#### 2.4.7 统计认识的信息符号

信息是客体内在的本质属性,当然客观信息系统与客体系统保持着同构。但是信息是统计认识主体对客观信息的认识、理解和把握,因而统计认识主体所获得的信息是否与客体的客观信息同构取决于主体的认识水平和能力。信息又是依赖于信息符号而存在的。信息符号是人们交往过程中,经过共同约定而成的用以指称、标志某一对象,传达某种信息的有意义的媒介物。所以信息符号是信息的指称和表征,是信息的载体和存在形式。来自客体的信息是信息符号的具体内容。因此,信息符号也必须与信息以及客体具有同构性。这是正确认识客体的前提。只有这样,统计认识主体才能通过信息符号的变化,客观地、准确地、观念地理解、把握和转译信息,进而完成对客体的认识。可见,更确切地说,统计认识是信息符号操作,是由信息符号将信息带入统计认识主体的认识图式,并使主体能够撇开实际客体而借助于信息符号系统来推演、再现和整合客体的本质属性和规律。信息符号将来自客体的信息带入统计认识主体的认识图式,实现客体信息的内化,主体通过信息符号的操作、推演观念地把握客体,然后再将认识成果表现出来实现外化,最后传播给社会使之转化为人类知识财产。完成了一次统计认识活动。可见,统计认识过程是通过信息符号内化→操作→外化→社会化来完成的。

统计信息符号作为统计认识的一个中介环节,在统计认识过程中具有其独特的功能。

标识客体,即表明统计认识客体的特征。这是信息符号最基本的功能。产生信息符号的动因就是人们为了记录、标志被研究客体的特征。人们在交往过程中,经过约定而形成了各种符号系统。如果一种符号不能明确指称某一对象或者特征,则

这种符号也就完全失去了意义和存在的价值。标识客体是符号的神圣使命。

载荷信息。符号的内容就是信息。符号是信息的存在物,信息是符号所携带的关于客体状态、属性、规律的反映。构建指称统计认识客体特征的符号系统,就是为了记录、传递、加工、存贮、传播客体的信息。信息就是客体状态特征及其在不同条件下的变化。它是客观存在而不实在。人们必须借以某种载体才能观念地把握和交流,才能实现信息操作。可见,在统计认识过程中认识主体的真实目的是为了把握客体的特性,但既不是通过实体操作,也不是信息操作,而是信息符号操作。这就是说,为了认识某一客体,人们除了通过对物理系统进行析取(采集)而构建的信息系统以外,客观上还存在一个符号系统。只有物理系统、构建的信息系统和符号系统三者保持同构才能保证认识结果的客观性。这种同构性越强,认识结果越客观、真实。事实上,随着信息处理手段和技术的发展,统计信息的加工和处理已经不再是多么困难的事了,统计认识的最困难、最关键的问题是如何保持上述三个系统的同构。

内化信息。简单地说就是将统计认识主体外部存在的关于客体的信息转化为主体思维活动的逻辑结构和认识图式。统计认识是统计认识主体通过思维进行的。一般说来,除了为了采集信息的目的,在实验过程中进行某些物理操作之外,统计认识一般都是在思维领域展开的。为了将客体的信息导入主体的思维活动之中,必须借助符号系统。离开信息符号的媒介,信息永远游离在思维活动之外,融于客体物理活动之中,实现主体和客体之间的交流。

加工信息。就是对导入的关于客体的信息进行整合、推演、获得新质的信息产品。信息是客观存在而不实在。因此,统计认识主体无法直接对信息进行操作,需要借以载体来实现。统计认识活动是统计认识主体对携带着来自客体信息的信息符号

的操作过程。这种信息符号操作包含有三层含义：一是信息以信息符号的形式存在并参与推演全过程。在统计认识活动中，信息符号往往表现为统计指标名称和指标值，统计认识活动中的每一个数据，都是特定时间条件和空间条件下客体某一特征的标识指代。二是推演方式方法规范为一系列公式符号系统。一种公式符号系统规定着一种特定的加工目的和加工方式、方法。三是加工过程是符号之间的运算。整个加工过程从形式上看就是符号之间的交流，就是公式符号系统与数据符号系统的内在结合。但实质上是符号载荷着信息进行的信息交流的转换。

外化信息。经过信息加工，统计认识主体对客体有了新的认识，获得了新的信息，主体将这些认识成果表达出来同样需要以信息符号为载体，才能使之从统计认识主体的思维活动中转化出来。

社会化信息。就是通过社会约定俗成的信息符号将统计认识主体的认识成果传播给全社会，转化为全人类共同的知识财富。

认识的中介环节构成了认识产生的参照背景。客体信息在这个参照系背景中被选择、变换、改造和建构。这样，主体通常把握客体信息都不可避免地不同程度上受到了某些扭曲、变态。因此，我们说统计认识是一个不断地在中介中建构着自身的过程，也就是说，我们所认识的信息虽然与客体本身存在着某种对象性关系，但是，到达于我们主体的、被我们把握着的客体信息本身，则只能是通过多级中介的多级建构而产生出来。所以说，统计认识不是已有信息的选择，而是一个新的信息样态的创生或建构。当然这种创生或建构中并不排斥信息选择过程的参与。

中介环节本身并不是认识过程中要把握的对象，似乎是一个信息的转运站和加工站，它们对客体的信息进行了某种自适

应特性的变换、选择和建构,然后又逐级传送到后续的中介环节,由此构成了一个逐级信息变换选择和建构的链条。在这链条的每一环节上,客体信息都将是一种重建后的再现。这个相对重建就发生在每一前后相继的环节的相互作用之中,而每一个这样的相互作用都必将使它输出给下一个环节的信息带有这两个环节相互作用的色彩,或叫做相互规定性。通过了诸多相互作用的相互规定性,客体的信息便在主体中被建构了出来。由于这个相互作用的链条的起点是从客体开始的,所以这个在主体中被建构出来的信息样态便仍然保持着与客体特性的某种对应性。

## 2.5 统计认识宿体

### 2.5.1 统计认识宿体的界定

统计认识宿体是指利用统计认识成果从事各种活动的个人和群体。简单地说,就是统计信息的使用者。随着科学技术的迅速发展和统计认识的系统化、专业化、精深化,一方面统计认识成了人类各个领域认识客体的主要方法。另一方面,又逐渐向统计信息产业化方向发展,形成了一个专门的活动领域和相应的专业人员队伍。这种发展导致了统计信息产品的生产者和使用者相分离,使用者勿须自己进行统计认识研究,需要时可以交由专门人员或机构来完成。生产者也不再是仅仅为了自己的需要才进行统计认识活动,可以根据社会的需要从事专门化的统计认识研究活动。这样统计信息产品的生产者形成了统计认识的主体,统计信息产品的使用者就是所谓的统计宿体。称之为宿体是从信息论的角度考察的,他们是统计信息产品的归宿点,故将这一群体称为统计认识宿体。

统计认识宿体与统计认识主体的划分就象生产者和消费者

的区别一样,是相对于统计信息产品的不同作用方式而言的。如果是生产者的角色,从事统计认识活动,就是统计认识主体。相反,如果是(消费者)使用者的角色,不从事具体的统计认识活动,则就是统计认识宿体。主体与宿体也有一致的时候。当完全是为了自己从事实践管理研究的需要,而自己进行统计认识研究,自己又是统计认识成果的直接使用者,这时,统计认识主体同时也是统计认识宿体。就象物质活动中自产自销者一样,自己既是生产者也是消费者。可见,统计认识主体和统计宿体的分离,是统计认识专业化、产业化的产物,并随着统计认识产业化发展,分离越来越明确。

统计认识宿体的形式也是多种多样的,也可分为社会宿体、群体宿体和个人宿体。社会宿体就是全社会共同使用某一统计认识成果。统计认识成果是人类共同的知识财富,属全人类享有,且可世代相传。从这个意义上说,社会宿体是统计宿体的重要方面。但是,也有相当多的统计认识活动是专门为某一机关、机构或个人而进行的,这种认识成果仅供、至少是首先供它们使用。因此,群体宿体和个人宿体也是大量地客观地存在着。

### 2.5.2 统计认识宿体的功能

提出统计认识需求。统计认识活动是为了满足一定宿体的具体需求而专门组织的,具有很强的针对性。没有统计认识需求,不仅不会有统计认识活动,而且统计认识主体、统计认识也不会存在和发展。统计认识需求来自于人类改造世界、管理社会和科学研究活动。在这三项活动中,宿体(想要使用统计信息产品的人们)提出需求,统计主体组织统计认识活动,如同有一定的市场需求,方才组织物质产品生产一样,这些需求者将是产品的使用者。可见统计认识宿体是统计认识活动的动力源泉。

使用统计认识成果。统计认识活动的重大意义主要体现在统计认识成果的使用面、使用程度及使用以后产生的效果等方

面。统计认识宿体就是统计认识成果的使用者,所以,使用认识成果这是宿体的天职。严格地说,不使用统计认识成果,它就不是统计认识宿体的组成部分。因此,在统计认识活动中不能将人简单地进行二分法,认为除了统计认识主体之外的所有人组成统计认识宿体这是不全面的。在统计认识宿体中有时可能也包含统计认识主体,同时,也存在既不是统计认识主体,也不是统计认识宿体的人群。统计认识宿体定义的核心词是“统计认识成果的使用者”,换句话说,是统计认识成果的使用者,就是统计认识宿体,统计认识宿体一定是统计认识成果的使用者。既充分,又必要。

检验统计认识成果和反馈信息。统计认识是否客观、准确、深刻,需要进行检验,检验的客观手段和标准就是实践,就是统计认识宿体使用统计认识成果的实践活动,统计认识宿体在使用统计认识成果从事的实践活动中,对统计认识成果的客观性、准确性、精确性及合目的性进行检验,并将检验结果反馈给统计认识主体进行更深入地认识研究。实现实践→认识→再实践→再认识的螺旋式循环,由不知到知,由知之甚少到知之甚多,由浅入深,由低级到高级,逐渐逼近客观真实。可见,没有统计认识宿体的实践活动,就没有统计认识主体的统计认识活动。统计认识宿体是统计认识系统的一个重要组元,是统计认识的动力源泉,既是统计认识活动的出发点又是归宿点。

# 3 统计方法论

## 3.1 统计方法一般

### 3.1.1 统计方法的普适性

#### 一、方法的层次性

“方法”一词是从希腊语“沿着”、“道路”而来的。因此,方法从词意上来看,意味着沿着(正确)道路运动。方法是人们根据某一领域的客观规律所制定的那些用来进一步认识、改造世界获得新成果所应用的方式和手段。人们认识、改造世界,必然要进行一系列的思维和实践活动,这些活动所采用的各种方式,统称为方法。根据布勒的定义,方法被规定为“一种在可以再现的程序范围内善于自觉地有计划地处理自然综合体的能力”(《方法概念》《The Concept of Method》纽约和伦敦,1961年,第135页)。按照这一定义,方法就是作用于被研究客体的一定程序,一整套手段。表现为把某种理性的

体系应用于主体的思维(理论)和实践活动过程中的种种现象。正如黑格尔所说的那样,方法“被列为一种工具,是某种站在主观方面的手段,通过它,主观方面和对象发生了联系”(黑格尔:《逻辑学》下册,商务印书馆,第532页)。由于方法是以各科学部门所形成的那些符合客观真理的知识体系为根据而形成的,因此,多种多样的知识体系产生了丰富多彩的科学方法。有些方法普遍适用于各门科学,有些方法则为某一些科学共同采用,有些方法则仅为某一科学所采用,还有一些方法仅仅为某一科学的非常狭窄的研究领域所采用。

不论是实践活动、信息活动、还是思维活动,在其任何一个领域中都有自己的不同于其它活动形式的方法。如化学方法不同于物理学方法,心理学方法不同于语言学方法,金属加工方法不同于计算机制造方法等等。从方法的普遍性程度、理论论证的性质和运用某些科学认识方法的领域的不同,科学认识方法论可以分为不同层次:

第一层次:认识中某个非常狭窄的领域所特有的和专门的、有时带有经验性质的各种技术手段、规定和作业,构成科学方法论的最低层次,它同操作方法相吻合。

第二层次:适应于某一个个别专门科学的领域(如化学、生物学、物理学、经济学、语言学等)使用的方法。这个层次上的方法的特点在于,使用有关的方法的自觉性不限于知道“怎样”和“在哪里”使用它们,而且还在于理解它们的理论基础。科学的方法依据相应的规律性的知识,也就是依据该对象领域的理论。理论带有方法的性质,满足方法论的功能,不仅成为解释现有和预测未来的武器,而且是探索和揭示现象的新的性质、更深规律性、更深刻的本质的手段。

第三层次:适应于在某些学科领域应用并以各个对象领域共同的规律性的存在为依据的方法。这些规律性的揭示,创立了相应的理论,反过来,这些理论又成为更加普遍的、跨学科方



法的理论基础。属于这一层次上的方法有:某些控制论方法,符号学方法,某些数学方法等。处于这个层次上的方法,虽然其应用领域超过了一门学科的范围,但整体上来说,仍不带有哲学性质。不过,这里已经发生了哲学问题,这些问题主要同某些基本概念的认识论分析相联系,同对于相应方法的认识论意义和适用界限的评价相联系,同阐明学科之间的相互关系相联系。

第四层次:适用于一切科学领域的方法。这是最普遍的科学认识方法,它们是哲学方法论研究的对象。当然,这并不意味着只有哲学方法论才研究和应当研究科学认识的普遍方法。普遍的科学方法,在反映每一个具体历史阶段上的科学知识状况时变化着、完善着,在科学技术进步的过程中,同自然科学、技术科学、社会科学以及逻辑和数学的成就一起变得日益复杂化,例如,公理方法是演绎方法的现代形式,在这种形式中反映了形式化理论体系的特点;现代归纳方法不仅吸收了关于存在和认识规律性的概率性质的观念,而且吸收了被专门加工过的数学概率论的技术。

## 二、统计方法的“普遍有效”性

统计方法就是指统计认识活动中所采用的一切方式和手段。统计认识从观测开始,由观测而获得描述事实的数据(逻辑形式上表现为单称陈述),然后经由归纳而发现客体的一般特征和规律性(在逻辑形式上表现为普遍陈述),即运用归纳方法从单称陈述推出普遍陈述,最后从总体一般特征、规律性和先行条件的合取中演绎出预见和解释,如果预见和解释被证实,统计规律性也就被确证了。如果需要的话,还要根据宿体的要求和客体的统计规律性并结合环境条件做出最优决策调控方案。前三个环节是每一次统计认识活动都必须完成的,而最后一个环节往往是为了统计认识研究的目的是为了控制管理时才是必要的。其基本关系简示如下图:

观测数据→统计规律性→预见和解释→决策调控方案

这种方法显然包含三个要点:

1. 统计认识开始于观测(包括实验中的观测)。观测中所获得的描述事实的原始信息是统计认识的基本素材,是统计认识的基本支撑点。

2. 归纳方法是从大量观测数据中发现统计规律性的基本方法。虽然归纳方法本身是一种推理,但是通过改进归纳程序,将分析、比较、排除等方法应用于归纳过程,可以克服简单列举归纳的不可靠性,使归纳推理符合科学推理的要求;当把概率概念引入归纳过程,形成概率推理时,就可以据此分析归纳推理的可靠程度,使之具有与演绎推理相比较的严密性。通过把统计引入归纳过程,可用严格的数学理论为归纳推理提供精确化的形式。上述种种改进,使统计归纳方法可以满足以下条件:形成概括基础的观测陈述的数目足够多;观测可以在各种各样的条件下予以重复;可以使推出的结论与公认的观测陈述没有冲突。

3. 统计认识成果的确认是在演绎结论与具体事实的对照中实现的。如果演绎结论在实验中被证实,那么统计认识成果本身就得到一次确认,被证实的结论越多,统计认识成果的确认程度也越高。虽然用以确认统计认识成果的具体事实必须与作出发现时的归纳证据有所不同,但这种确认过程原则上依然符合归纳程序。

统计认识方法一开始在认识运动中所起的作用是极其微小的。但随着认知体系(实质性科学、哲学、技术等科学理论和认识水平)的发展,统计方法不断向科学、深化、精确发展。统计方法是辩证的、发展的。现在任何一门科学都非有这些方法不可。由于统计方法不属于哲学方法,(因为哲学方法试图揭示人类思维向真理运动的规律,拥有发达的、内容丰富的范畴体系作为自己的逻辑武器,是作为所有其他方法的总括而产生的),但是,它是各门科学都应用的方法。为了区别起见,称之为专门的通用科学的方法。也就是说,从方法据以形成的理论体系(概念、规

律)来说,统计方法不属于哲学方法而属于专门方法;从方法应用的范围来看,统计方法不是仅仅为某一门科学所采用,而是有普遍适用性即“普遍有效”性。在这一点上,具备哲学方法的普遍有效性的特点。因此,至少可以说作为关于统计方法的学说统计方法论是一门“亚哲学”。事实上,统计认识方法的显著特点就是根据客体的信息不断获得认识客体属性的一个认识过程,通过不断修正,使统计认识从广度和深度两个方面无限逼近客体的本身属性。在这个意义上说,统计认识方法又揭示了人类思维向真理运动的规律,规定认识向各种真理运动的一般途径,具有哲学方法的属性。此外,从哲学方法的构成要素来看,弗·培根提出的方法的组成要素为:归纳、分析、比较、观测和实验。统计方法也是以此为构成要素的。统计学建筑在观测、实验和分析的基础之上,同时又上升到认识原因、规律和要素之间的关系;统计学家经过统计认识实践制定了采集数量信息的方法和理性加工数量信息的方法,其突出特点是把观测与实验同精确的数学分析和数量表示经验所得的成果结合在一起。

### 三、统计方法的基本性质

统计方法是统计认识中所用各种方法的总称。作为统计领域的基本要素之一,各个领域的统计认识研究都有自己特有的方法,而且在统计认识的不同环节也有着不同的统计方法。但它们有着许多共有的特性:

1. 统计方法应当是可操作的。主客体的相互作用是通过某种操作来实现的。在认识系统中,一定的方法只有与一定的操作过程相联系,才能把系统的各个要素连接起来。原理、理论一旦可以操作,就转化为方法,从而成为行动中的理论,或在主客体相互作用中起实际作用的理论。

2. 统计方法应当是有规则的。统计方法不是盲目的,无条理而是要遵循一定的规则和程序。当然,规则并非单一的、固定不变的,也不一定都能形式化为逻辑格式。

3. 统计方法应当与研究对象相一致。统计方法的形成不是一蹴而就,而是要经过一个主客体的相互适应的过程,才能形成一定的方法。在这个过程中形成的方法,必须基于对象的本质并在对象本性所允许的范围内发挥作用。方法若与对象不一致,方法就会失去有效性。

4. 统计方法应当是有结构的。统计认识是在多种方法交替作用过程中进行的,每种方法均有不可替代的作用,而方法之间是相互关联、渗透和耦合的,从而成为结构的整体。

### 3.1.2 统计认识方法的主观方面和客观方面

统计认识方法是统计主体实现统计认识目的、完成统计认识任务的手段。解决“怎么样”才能借助客体信息正确把握客体的问题。它是实现主客体之间信息交流与建构的中介。毛泽东说:“我们不但要提出任务,而且要解决完成任务的方法问题。我们的任务是过河,但是没有桥或没有船就不能过。不解决桥和船的问题,过河就是一句空话。不解决方法问题,任务也只是瞎说一顿。”(《毛泽东选集》,一卷本,人民出版社,1964年4月第1版,第125页)。毛泽东用“桥”和“船”在过河中的作用来说明方法的重要性。黑格尔也曾说过:“手段是一个比外在合目的性的有限目的更高的东西——犁是比犁所造成的、作为目的直接的享受更尊贵一些(黑格尔:《罗辑学》下卷,商务印书馆,第438页,)。”从某种意义上说,一种正确的方法,其价值比由它获得的直接后果更大。统计方法作为使主客体信息由此岸到达彼岸的工具,其作用和重要性无异于“桥”、“船”和“犁”,但是从其存在形态来看,它又不是“桥”、“船”和“犁”,它是统计理论工作者和实际工作者创造性运用科学理论知识的产物,是主观能动性的产物,没有创造能力,没有发挥创造性思维巧妙地运用理论的技巧、艺术,就不能从理论知识体系中提炼出方法来。也就是说,不能把理论变为行为活动的指南。

因此,在表面上看来,统计认识方法表现为某种主观性的东西,与认识客体相对立。认识主体借助于一定方式的认识体系作用于被认识对象来实现自己的认识目的。如果把这一方面过分夸大,使之僵化,就会把统计方法视作是与客体格格不入的一种纯主观的程序。但是,果真这样的话,统计方法也就无法引导人们去把握客体。统计方法是一种认识方法,在认识过程中,统计方法作用于数据(即信息),是一种信息操作,是认识主体对客体的一种观念把握。它是根据客体的信息运动规律和数量规律性以及质与量的辩证规律而制定的采集、加工、存储、开发、提供、利用信息,以实现认识和调控客体的一系列行为规则。

正确理解统计方法必须客观地把握两个方面:客观方面和主观方面。所谓统计方法的客观方面是指统计认识方法本身是一种客观真实的知识体系,是对信息(更确切地说是数量信息)运动规律、数量规律、质量互变等客观规律的认识,或者说客观规律被“移植”和“被移入”到人类意识之中,转化为认识主体的行动规则和认识手段。是已经被认识了规律性。它决不是天赋的,也不是主观任意构造的。统计认识方法的主观方面是指统计认识方法是人们为了解决“认识主体应当怎样做才能获得更客观、更合适、更准确、更深刻的认识”这一问题而研究制定出来的,是认识主体为了实现某种目的而研究和选择运用的,是主体连结主观与客体的工具。可见,统计认识方法的客观方面是统计方法的生命,是本质内容,而统计认识方法的主观方面仅仅是一种表面形式。但是,恰恰就是这种表面形式常常被夸大,绝对化,因而把统计方法误解为与被认识客体无关的工具,是“纯粹”的方法、绝对的方法,主体可以根据个人意志主观地选择运用。同样,过分夸大统计方法的客观方面,而否认其主观方面也就无方法可言了。所以,二者相辅相成,缺一不可。

统计认识方法的客观性决定了统计认识主体在选择运用统计方法时,必须根据客体的具体内容和性质来选择,要使方法与

客体的规律性相符合。任何统计方法都有其一定的应用范围,这个应用范围就是被认识了规律性,就是该种方法的客观方面,超出这个范围以外去应用这种方法就不可能得到正确的结果。因此,在确认统计认识结果时,必须考察统计认识活动过程中原始统计信息的代表性、统计方法运用的正确性和统计认识结果的真实性。代表性是指籍以进行统计认识研究的原始信息是否比较好地反映了认识客体的基本性质。是原始材料与客体状态间的比较,评价的是原始信息对客体的记录、反映程度,是建构的原始信息系统与客体系统的同一性问题。这种代表性首先表现在原始材料来自于客体;其次原始信息的采集要科学合理,既没有信息遗漏,又没有信息冗余;再次原始信息不含噪声,起码噪声甚少,虽然有失真,但失真度不大,信噪比较高。正确性是指统计认识活动中正确地应用了各种统计方法。是统计活动过程与客体状态之比较,评价统计认识主体的行为是否遵循规则,统计方法是否符合客体的规律性。是统计方法与客体的同一。换句话说,评价的不是认识成果、思维内容,而是主体的行为。真实性是指统计认识成果客观、真实地再现了客体的属性。是思维内容与客体的比较,评价的是认识成果与客体的相符性,是认识成果与客体属性的同一。

代表性、正确性和真实性三者是密切联系的,原始信息的代表性是基础,方法应用的正确性是关键,认识成果的真实性是目的。事实上,原始信息的代表性也取决于采集方法的正确性。从这个意义上讲,统计方法的正确性是整个统计认识活动的灵魂。信息采集方法的正确性,可以使我们获得有代表性的原始材料,在此基础上统计分析方法的正确性,生产出了具有真实性的统计认识成果。

统计方法的正确性不在于统计方法本身,而在于统计认识主体的具体运用。因为任一统计方法和其他一切科学方法一样,都是一定理论体系的具体化,也就是说方法是工具化了的知

识体系。因此,只要在这个理论体系范围内,统计方法本身不可能是不正确的,不正确是由于统计认识主体在超出了产生该种方法的理论体系以外使用了这种方法而造成的,或者说是统计方法的主观方面的不正确性。

黑格尔曾经揭示了方法的客观依据,即一种真实知识的体系反映了对客观规律性的认识。这些规律性变换成了主体的行动规则。正如 T·巴甫洛夫所指出的:“科学方法乃是作为客观世界主观反映的人类思维运动的内部规律性;或者也可以说它是‘被移植’和‘被移入’到人类意识中的客观规律性,是被用来自觉地有计划地解释和改变世界的工具。”([苏]柯普宁著,彭漪涟等译:《作为认识论和逻辑的辩证法》,华东师范大学出版社,1984 年版)

在统计认识方法中,方法的客观方面即客观规律性转化为统计认识主体的行动规则。因此,任何统计方法都表现为一整套为统计认识而制定出的规则或手段。比如,抽样方法表现为如何制作抽样框,怎样实施随机取样,抽取多少样本单位,如何计算样本指标及其抽样平均误差等一系列行为规则。这样就存在一个认识主体行为的正确与否的问题。所以通常被称为统计认识方法的正确性。可见统计认识方法的正确性与真实性是两个含义根本不同的概念,二者不能相互取代。统计认识方法的正确性是指认识主体行为的合规则性,即通过将认识活动与规则手段进行比较,评价认识主体的行动是否严格遵循方法的规则。(正确性借助于据以制定出行为规则的知识体系的真实性而与客体相关)。而统计认识方法的真实性是指思维(认识)内容与客体的同一性,即通过将思维内容与客体进行比较,评价思维内容与客体是否相符。

### 3.1.3 统计认识方法的突出特点

#### 一、定性与定量认识紧密结合

统计认识是通过量的分析来达到对客体质规定性和量的确定性的把握。统计方法本身不论其数理性怎样,但是,获得什么样的信息、怎样获得、模型解读等都需要一定的定性分析。

统计认识是从设计统计指标开始的。从认识论来看,设计统计指标实质上是统计主体观念地建构客体的过程。不论是认识何种客体,统计认识主体必须首先解决从哪些方面来把握客体。即必须首先研究客体的主要属性、结构、功能及其表现。而要解决这一问题,必须借助于被认识客体的相关实质性科学的概念、范畴、规律等基本理论进行定性分析。在此基础上,经过量化、综合、形成统计认识指标体系。统计指标体系实质上就是客体系统的结构框架的观念形式,我们称之为客体系统的观念系统。

搜集统计信息也是定性与定量相结合的一个过程。在搜集统计信息时,不仅要根据客体的特殊性质选择适应的信息搜集方法,而且在获取信息的具体过程中,必须确认信源与统计指标的规定同一。就信息搜集方法而言,对于可控客体来说,一般适宜于运用实验方法;而对于不便于控制的客体则宜用观测的方法;有些客体可以全面观测,有的则宜用非全面观测;有的适宜于直接获得信息,有的则宜于运用次级信息(例如,对于一些被调查者比较敏感、不愿意、或不便于直面回答的信息,往往根据相关资料进行计算和推算的结果更客观。)等等,所有这些问题都需要在对客体进行定性分析的基础上加以解决。就信源与指标的同一起来说,统计信息搜集工作者必须确保任一数据都完全符合统计指标的规定。这是获得有效的统计认识成果的基础。比如,指标的范围、口径、计量单位、实验条件、测量手段、计算方法等必须与指标要求相符合。这一阶段实质上是同构客体系统的信息系统,是使指标体系建构起来的结构框架具体化,使之有血有肉、有活力,是实体系统的再现。接下来的数据处理和认识研究都是以这个同构的信息系统为对象进行的。可见,只有原



始信息与客体同构,统计认识成果才有可能与客体同一。同构(Isomorph)就是指两个系统在功能上是等效的,而在表示方法上不同。地图与它所表现的地域同构,底片和由它冲晒出来的照片同构;还有格局方面的同构。

统计信息加工整理的核心内容,一是确认原始信息的客观性、真实性,即与客体的同一和与指标体系的同一。二是显示客体的分布特征。在揭示和显示分布特征时,统计分组是否科学十分重要。从理论上说,分组应是客体质变的再现。每一个组就是客体一定层次的一个度。度是客体的质和量的统一,是客体保持稳定的“量变范围”,也就是客体的质所能容纳的数量关系。组限就是该度的临界点。因此,这两个核心内容都是定性分析和定量分析的统一体。

统计分析也是一个定性分析与定量分析相结合的过程。这一环节往往被人误解为一个定量加工阶段,特别是计算机的运用,似乎完全是一个数字计算过程。其实不然。统计方法本身不论其数理性怎样,但是,获得什么样的信息、怎样获得、以及模型解读等都需要一定的定性分析。数字计算是按照一定的统计方法计算的,而任一方法都有其应用条件,也就是基本假设。因此,统计分析问题归结为选择统计分析方法问题,而选择统计分析方法问题归结为检验方法假设与客体属性的相符合问题。所以,看起来是一个计算机的数字运算过程,但实质上这一过程的关键点在于分析统计方法的应用条件与客体属性特征的相合问题。不论是什么样的方法,计算机都可以进行计算,并给出一个结果。但是,如果违背了方法的基本假设,方法的应用是不成功的,得出的结果也是没有根据的。

统计解释、评价和预测趋势、制定决策方案更是定性定量分析的结合体,这已不言而喻,在此不再重述。

## 二、统计方法是归纳推理与演绎推理的辩证统一

从经验出发形成一般原理,而又从一般原理推出新的特殊

事实。用新的观测检验这些特殊事实,这就证明以前形成的一般原理是正确的。这是伽利略的研究方法。

达·芬奇把获得真理知识化分为两种要素,即经验和数学。关于真理的科学建立在慎重提出并检验过的经验和深思熟虑的观测基础上,该科学借助于与真理相符的结论从经验进到对规律性的认识。认为量的数学研究方法是一种普遍的方法。达·芬奇在《达芬奇选集》两卷本(转引自《作为认识论和逻辑的辩证法》,彭漪涟、王天厚等译,华东师范大学出版社,1984年版,第192页)中指出:“如果科学不能应用任何一门数学,不与数学发生联系,那科学就一点也不可靠。”

培根的推理学说的重要之点在于强调观测和实验在推理中的巨大作用;归纳法应该依靠经过仔细研究、检验并整理成一定秩序的大量事实。经验事实是推理的首要 and 重要的前提;在进行归纳推理时,不要急于用区分法和排斥法加以概括。

笛卡尔则承认直觉和演绎的决定作用,把它们看作是认识一切的两条最可靠的途径,并在此基础上建立起自己的认识理论。只有直觉和演绎才可靠,其余的一切都可怀疑并将导致谬误。同时,直觉比演绎更可靠。

统计作为归纳方法,获得结论按其性质来说还是有问题的,需要予以检验。现实生活的实践和科学的发展,使凭借归纳途径获得的结论不断精确和改变,只有通过归纳、演绎和实践的相互作用才能找到可靠的科学真理。归纳是由一般性程度比较小的知识推出一般性程度比较大的知识的推论,而演绎正好相反。

### 三、可靠性与创造性相结合

在人类认识中普遍存在的一个矛盾就是认识的可靠性和创造性的对立。不论是认识主体,还是认识客体,一方面都希望认识结果尽可能地准确、可靠,另一方面又希望有一个新的发现,对客体有一个更深入的了解和把握。我们知道,要得到绝对可靠的认识必须运用演绎推理,并且推理所依据的大前提和小前

提要保证正确无误。这就客观上决定了得到的结论不可能超出前提的范围,也就是说,认识结果不会有任何创造性。而归纳推理充分发挥了认识主体的创造力,认识结果是一个新的发现,但又无法给出认识结果的可靠性。统计认识是在一定概率提供的可靠性下,充分发挥创造性而进行的探索,其认识成果是有可靠性保证的新发现。并且,统计认识主体可以根据不同的认识目的、不同的认识客体等具体问题,相应地调整认识成果的可信度和创新程度。使统计认识的可靠性和创新性统一于一体。

#### 四、分析与综合统一

所谓分析方法,就是把研究对象分解为若干组成部分,并分别加以研究,从而认识事物的基础或本质的一种思维方法。

任何事物的整体都是由若干组成部分构成的,将客观事物在一定条件下分解成各组成部分,分别研究其结构与功能、各部分相互联系、相互作用的特点以及在各种外界条件作用下所表现出来的事物的各属性和特点,从而达到对事物本质及内在规律性的认识之目的。可见,分析方法是客观事物的整体与部分关系为客观基础的。在运用统计分析方法时,主要涉及到以下三个方面的客观事物整体与部分的关系:一是分析事物在空间分布上整体的各个组成部分的分配和布局;二是分析事物在时间发展上整个过程的各个阶段的状态及其走势;三是分析复杂的统一体的各种因素、方面和关系。

从思维方式上看,统计分析方法最突出的特点是从事物的整体深入到各组成部分,通过深入地认识事物的各个组成部分来认识事物的内在本质或整体规律。使认识从一个层次发展到更深的层次,从现象的认识深入到本质的认识。因而,统计分析方法在科学认识发展中具有重要的意义,几乎贯穿于统计认识的全过程,并渗透到其他研究方法之中。

当然,统计分析方法也有其局限性。对事物进行必要的分割、孤立的分析研究,虽然能把人的认识引向深入,但也可能将

人的眼光限制在片面、狭窄的领域里。人们认识事物,不仅要认识它的部分,更要认识它的整体,这就要运用综合方法。此外,在进行统计分析时,有时往往是在暂时割裂、孤立的状态下,并假定除被研究因素以外其它因素都不发生变化的条件下进行研究的,这是不符合客观实际的,因而统计分析结果也是不十分精确的。

所谓综合方法,就是把研究对象的各个部分联系起来加以研究,从而在整体上把握事物的本质和规律的一种思维方法。与分析方法相比,综合方法认识过程的方向完全相反。它是将事物的各个部分联结为整体,通过全面掌握事物各部分、各方面的特点以及它们之间的内在联系,并加以概括和上升。从事物各部分及其属性、关系的真实联结和本来面目,复现事物的整体,综合为多样性的统一体。但是,这种综合绝不是把事物的各个部分机械地连在一起,而是按照事物的内在联系,在思维中把各个部分结合成一个有机联系的整体。同时,综合也不是简单地回复到对整体的混沌认识,而是在分析事物细节的基础上,揭示事物的本质和规律。

许多科学假说、科学定律、定理、公式的提出,以至科学理论体系的建立,都常常需要综合人们对事物各个方面的认识,综合许多人的研究成果。随着现代科学发展正日益出现整体化趋势,学科高度分化而又高度综合,只有运用综合法,才能更好地开拓新领域,建立新学科。一系列综合性学科、边缘学科、横断学科的出现,无疑是与综合方法的运用密切相关的。

从辩证法来看,统计认识不在于把整体分解为部分,从而揭示出这种整体由它们所构成的那些最小单位,而在于证明这个整体由那些要素、并以何种方式产生和发展起来的,以及这些要素在整体的形成发展中起着什么作用。例如,不论是静态模型,还是动态模型,都是先将客体分解为若干个主要因素,然后又将其结合起来考察其综合作用。

统计方法不是所得知识的某种标准、检验等级,而是通过批判分析具体客体的数据信息增加实际知识的一种工具、方式和方法,是具体分析现实对象,现实事实的一种方法(方式)。应当指出,综合法总是与分析法并用的。没有分析的综合,其结论就只能是空洞的、无根据的,是一个混沌的、外在的、直观的整体。“思维既把相互联系的要素联合为一个统一体,同样也把意识的对象分解为它的要素。没有分析就没有综合(《马克思恩格斯选集》第三卷 人民出版社 1972 年版 第 81 页)。”分析的结果,也就是综合的出发点。统计认识的发展总是沿着“分析——综合——新的分析——新的综合……”轨迹不断前进的。

## 3.2 客体的信息建构

### 3.2.1 客体的观念建构

统计认识是以客体的信息系统为媒介的,而信息又是客体存在方式、运动状态以及各种属性的表征,因此,信息离不开客体特征的“名”即概念。显然,在统计认识中欲建构客体的信息系统,必须首先建立相应的概念系统。

统计指标是概念在统计认识方法中的具体化。概念是整个认识思维的一个基本形式,它概括了客体的本质特征。在统计认识中,首先要从现有的科学知识和以往积累起来的认识、经验以及业已形成的概念出发,将概念具体化为一个或一组可观测的统计指标,并据此搜集大量的客观信息,然后借助于统计方法,按照客体的内在联系对搜集到的信息进行加工、开发,作出判断、推理和论证,获得新的认识,形成新的概念。所以,就统计认识过程而言,概念既是统计认识的出发点,也是统计认识的归宿点,整个认识过程是概念的运动过程,是概念的形成和发展过程。所谓客体的观念建构就是客体的概念系统的建构,就是统

计指标体系的建构。

统计指标是反映统计认识客体状态、属性及其运动规律的特殊形式。统计指标从形式上看是一些数字,但从其内容来说是客观的、有具体的内容,甚至最抽象的指标在统计认识客体中都有其类似物、原型或成份。指标反映的内容就是客体中所包含的东西。统计运算中的不可加性就是这一实质的具体体现。如2个人与3支枪是不能相加的,同样两架飞机与三只钮扣也是不能相加的,不能相加不是因为数字“2”与“3”本身,而是因为其具体内容不具有同质性,从而不具有可加性。

统计指标不是统计单位某一标志的简单记录,而是总体的本质属性、方面,是这些属性和方面的联系、关系及其运动、发展的规律,是总体普遍本性的反映。但这决不是说,指标中没有一般与个别的任何联系。恰恰相反,指标反映的是与个别相联系的普遍。为了(得到)形成指标,必须从研究大量的个别着手。恩格斯说:“人们曾用来学习计数,从而用来作第一次算术运算的十个指头,可以是任何别的东西,但是总不是悟性的自由创造物。为了计数,不仅要有可以计数的对象,而且还要有一种在考察对象时撇开对象的其他一切特征而仅仅顾到数目的能力,而这种能力是长期的以经验为依据的历史发展的结果。”(《马克思恩格斯全集》第20卷,人民出版社1965年9月第11版,第41页。)

统计认识同一切事物发展过程一样,是一个辩证过程,因而,在其思维和认识过程中必须要运用辩证唯物主义的普遍原则、规律和范畴。统计指标作为统计认识理论的思维形式,其形成过程首先要遵循辩证逻辑。具体地,一切统计指标都处于相互联系和发展之中,都要受到量变到质变、对立统一、否定之否定等规律所制约。

统计指标的形成一般有两种途径,一是已有概念的指标化,一是直接对原始事实进行理论概括。概念往往比较抽象、可测

性较低,而统计指标则必须具体、并且要有较强的可测性。将现有概念具体化为统计指标时,指标名称要尽可能与概念保持一致,最起码要有一种能唤起联想的关系,以便于宿体理解和掌握。对于复杂概念来说,必要时,可将某一概念具体化为一组统计指标,从不同侧面分别进行测量,然后综合起来从整体上全面反映概念的内容。有了指标名称以后,还要对指标的计算范围、计量单位、计算方法以及空间、时间范围等作出具体规定,以便于操作。在对原始事实进行理论概括时,需要通过复杂的抽象化思维方法来探寻原始事实本身所蕴含的必然的内在联系,从大量事实、事件或许多情境中析取或抽象出它们共同具有的性质或标志,经过比较、分析、综合等一系列理论思维和逻辑概括形成对客体认识的概念,然后再对所形成的概念具体化为统计指标。

认识任何客体的统计指标都是有层次的。有描述客体规模和水平的外延指标,也有表征客体某一性质的程度、速度、内部结构和外部关系的内涵指标。从统计指标的认识功能来看,有描述指标、预测指标、预警指标和决策指标。

某一个具体统计指标反映着客体的某一方面的特征。而任何一个统计客体都有许多特征,这就决定了一个被认识客体客观存在着一系列统计指标。并且各个统计指标都不是独立存在的,在一定范围或条件下内在地处于相互联系和相互制约之中。若干个具有各种相互联系的统计指标所形成的有机整体就称为统计指标体系,也就是客体的观念系统。

从人们的认识角度出发,人们对客体的认识往往不仅仅在于认识其某一方面,而是要了解和掌握客体的整体特征和规律性。由于单个统计指标所反映的毕竟只是所研究客体的现象或过程的某一方面或某个特征,是客体发展过程中某一时间上的抽象,因此,就其本质而言,它使某一事实片面化或单一化。为了克服这一不足,把认识客体作为一个整体进行系统地认识,就

不能只根据某一个统计指标对客体作出评价和判断,必须使用一套科学的、相互联系、相互依存、相互制约的指标体系,才能很好地完成统计认识的功能。如果不考虑或忽略各种统计指标与其它指标的联系,根据某个孤立选取出来的指标是无法达到对所研究客体进行深刻认识的目的。

一个统计指标只有从属于一定的体系之中,才有具体意义。统计中不存在某种绝对的、一成不变的统计指标。某一指标在各种不同的指标体系中可以从各种不同的角度进行计算,可以有不同的计算方法和计量单位等。

统计指标体系所反映的内容决不等于指标体系中各个指标所反映的内容的简单总和。指标体系的形成,从表面上看,似乎是把若干个有联系的指标排列在一起,多反映了几个方面的特征。实质上,任何一个指标体系都是一种新质的观念形成。统计指标体系除了包含各个指标的信息以外,它能提供关于各个指标之间相互联系的信息即各个方面相互制约、相互作用而产生的新信息,以及整个体系发展过程方面的信息。因此,统计指标体系永远比各个指标的简单综合更为重要。客体的本质寄寓于统计指标体系所反映的各个侧面的总和及其相互之间的关系之中。显然,统计指标体系的形成不能是把多个统计指标简单、机械叠加在一起,而是一个辩证的有机整体,各指标以一定严谨的逻辑体系共存于同一指标体系中,以全面、综合地反映和说明事物的本质。正因为如此,在统计指标体系中,指标之间的分布不是随意的。对任一事物的发展变化而言,其不同的影响因素,作用方式和影响程度均有所不同,使得不同的指标在指标体系中表现出层次性特征。有的居于主导的、基本的地位,有的则居于次要的、辅助性的地位。因而,科学、合理地建立统计指标体系,基本指标的确定至关重要。

客体观念系统的建构不是一蹴而就的,而是一个辩证的发展过程。这首先是由事物发展的必然性决定的。既然一切事物



都处于经常的变化和发展之中,那么,统计指标也必然随之而经常不断地补充、修订和更新。可以说,统计指标的发展是通过形成新指标和改进旧指标来实现的。新指标的形成反映了统计指标发展的突变性特征,是随着认识客体在其发展过程中新事物、新矛盾的不断出现而产生的。对旧指标的改进,则表现为统计指标发展的渐进性特征,反映人们对客体认识的不断深化、日臻完善的过程。其次,是由统计认识的发展性决定的。统计认识过程是不断深化、不断前进的。从不认识到认识,从对初级本质的认识到对高级本质的认识,表现为认识中的否定之否定。在统计指标的发展过程中,随着人类认识世界、改造世界活动的发展变化,统计认识也随之而发展变化,用以进行认识的思维方式——统计指标也将为适应这种认识上的不断发展变化的需要而或不断改进旧指标,或不断形成新指标,从而实现一次否定、继而否定之否定,……。

通过上述分析,不难得出如下观点:

1. 统计指标是第二性的。统计指标同其他科学概念一样,是第二性的,是通过人们的抽象思维形成的。统计指标具有客观的内容,但它并不等同于客观存在,并不是社会生活本身,而是社会生活的反映,属于社会意识范畴。人们经过抽象思维形成了关于物质存在的概念,以反映被研究现象的数量方面,这就是统计指标。

2. 统计指标体系具有客观存在的必然性。根据相互联系的原理,一切事物都处于相互联系之中。这种相互联系的形式可能是多种多样的:有本质的和非本质的联系,必然的和偶然的联系,经常性的和暂时的联系,函数的和相关的联系,因果关系等等。在运用统计指标解释和分析客体时,要尽可能地全面考察所研究的对象,即要从事物与事物之间、因素与因素之间客观存在的各种联系中去观察、分析现象的本质特征及规律性。而任何一个单一的统计指标只能反映事物的某一个方面的特征,因

此,对某一研究对象进行全面的观察和分析,就必须通过统计指标的科学体系来完成。

3. 统计指标具有时间和空间两种纬度。根据发展的原理,一切事物都是在不断运动和不断发展的,同时发展又是在时间和空间中进行的。因此,运用统计指标分析所研究的对象时,不仅要有静态指标以反映事物在一定时间、地点条件下的数量特征,而且要有动态指标以反映事物在不同时间、地点条件下发生、发展变化动态特征。任一统计指标所反映的内容都是客体在一定时间和空间条件下的状态和属性。

4. 统计指标是质与量的统一体。任何事物都是质与量的有机统一体。质是现象内部所固有的规定性或本质,量是现象的外部规定性,这种外部规定性往往作为现象某种属性的数值、数字和表现程度而显现出来。统计指标是社会经济现象质与量的统一。虽然统计指标是作为认识社会经济现象数量方面的手段,即以数字来表现事物的数量方面,然而,指标本身却具有其质与量两个方面。统计指标质的方面是指它的一般内容即概念上的特征而不涉及具体数量,通过反映事物质的特征将一事物与它事物区别开来;指标量的方面则是指它在时间上和空间上的界限。统计指标的质与量两个方面是不可分割的统一体,没有任何一种只具有孤立的质的方面的指标,也没有任何一种只具有孤立的量的方面的指标。按照质量互变规律,社会现象的发展是通过量的逐渐积累及其飞跃或渐进地转化为新质而实现的。

统计指标从现象到本质、从低级本质到高级本质的发展过程,就是思维从一种质转化为另一种质、一种更深刻、内容更丰富的质。统计认识的过程往往表现为由反映个别的个体指标向反映一般和普通的综合指标的转化,这种转化就是不断提高统计指标的理论概括性,是一种从个别到一般、从感觉到思维、从不甚复杂的概念到更为复杂的概念的一种“飞跃”。

### 3.2.2 统计观测方法

统计观测是获取原始信息的一种基本方法。它是统计主体通过感觉器官或借助于一定的观察、测量仪器,按照统计法则有目的地感知客体,从而获得原始信息的一种统计方法。所谓“原始信息”,是指统计主体对观测和实验等实践所观测到的客观存在的事件、现象和过程的信息。它是统计认识的基础。它所描述的或是独立于统计主体的意识之外而客观存在的事件、现象、过程,或是已经成为人们科学探索的对象、现象和过程,或是虽尚未有理论支持但已经由经验检验过的人们对客观事件、现象和过程的正确反映。从根本上来说,这些原始信息来自人们的生产实践和观测、实验等社会实践活动,同时还来自于利用前人和同时代人的研究成果以及通过实地调查研究从群众中获得的感性经验材料。巴甫洛夫曾经说过:“在科学中要学会做笨重的工作,研究、比较、积累事实。不管鸟的翅膀怎样完善,它任何时候也不能不依赖空气飞向高空。事实,就是科学的空气,没有它,你任何时候不可能飞起,没有它,你的理论就是无用的挣扎。”统计学是获得知识、规律的重要科学方法,是原始信息系统化的非常重要的形式之一。特别是在大量事件的领域内,统计学允许把大量经验材料整理成统计表格、频率分布、离散和相关,以及其它复杂的概括性的指标,使可能揭示所研究的事实的结构和规律性,并使对它们的理论理解迈出重要的一步。在统计认识过程中,任何统计指标、统计模型、统计假设以及统计规律都必须以一定的原始统计信息为依据。而且一切统计认识成果的真伪也必须经过来自客体的原始信息或具体实践的检验。因此,获得与客体同构的原始信息是进行统计认识研究的一个非常重要的基础环节。

统计观测是统计认识中第一性的和基本的认识过程。是统计认识过程中有目的的实践活动。由此我们得到关于外部世界

的第一手的、原始的信息。建构起客体的信息系统,实现由客体物理系统向信息系统转化。

统计观测方法有各种不同的类型。统计主体可以通过自己的感觉器官去观测、感知客体;也可以借助于各种科学仪器或依靠除统计主体以外的其他人群去观测客体,获得观测资料。前者为直接观测,后者则为间接观测。就观测的内容而言,既可以关于客体的质的特征,也可以是量的特征,由此观测方法可区分为质的观测和量的观测。不论是哪一种观测,其显著的特点是在获取原始信息的过程中,统计主体对观测对象不加任何变革和控制的自然情况下进行的。以此而区别于统计实验。

无论通过什么方式进行观测,最终都要从观测中获得有关观测对象各种特性的资料。这些观测资料是观测者通过自己的感觉器官或借助于科学仪器对观测对象经过大量的、反复的观测、感知,并将观测的结果用语言、文字或图像等方式描述出来而形成的。它虽然只是对观测对象的感性认识材料,但却是第一手的原始资料,具有极为重要的科学价值。任何统计信息加工活动都是以此为起点、为源泉的。正所谓“科学始于观测”。

统计观测虽然是感性认识的形式,但又不能归结为感性认识。它是对外部世界的对象和现象的有目的、有组织的知觉。它们是直接指向统计认识客体的基本的和原始的认识活动形式。是统计认识主体深入到理论层次把握认识客体的前提条件;而且是形成、检验和发展统计认识成果的实践基础。统计观测方法是基本的实践性方法。是统计认识中在经验层次上直接获得感性材料和原始信息的根本手段。同简单的感性知觉相比,以组织性、计量性和目的性为特征;同实验相比,没有对研究对象进行控制。法国学者 K 贝尔纳说过:“观测是在我们不能支配的自然条件下进行的”观测者“使用简单的或复杂的研究方法,对他所不能改变而只能收集的现象进行研究,也就是在大自然给他提供的这种形态下进行研究”(转引自《科学认识的方法

论问题》,知识出版社)

现代统计认识中的观测,固然是感性认识活动,却有着理性活动的特点。即要做到科学的观测,理性思维在其中起着重要的指导作用。统计主体不是消极被动地接受观测信息,而是在一定理论思维指导下,有目的地积极主动地探索问题。

统计观测还必须遵循辩证唯物主义认识论所规定的方法论原则——可观测性原理。即统计认识的范围应该是原则上可观测到的物质客体,也就是说,统计认识的对象可以被观测者的感官直接感知或者借助于观测仪器,间接地被观测者所反映、所把握。按照这个原则,各种臆想的本质(“心灵”、“灵感”等)以及可能和其它客体产生相互作用,因而不可能传递关于本身的任何信息的客体都应当排除在科学认识的范围之外。

观测仪器作为观测者和观测对象相互作用的中间环节,在观测中起着代替或加强人的感官的作用。从认识功能上讲,观测仪器乃是人的感觉器官的延伸。人类原有的感觉器官在接收和传递信息上有很大的局限性。但是随着科学仪器的发明和使用,这种生理条件造成的“认识界限”一再被冲破。观测仪器可以帮助人们不断扩大可观测的时空尺度,扩大感觉的范围,从而拓宽人的科学认识的视野。借助于科学仪器观测,人的感官无法感知的自然信息可被感知,曾几何时苦于分辨不清的观测变得清晰,帮助人们把本来不能直接接收和传递的信息转换成可以接收和传递的信息,从而为科学认识提供越来越多的素材。观测仪器可以克服人的感官在生理阈值上的局限性。观测仪器还可以克服人的感觉器官在接收和传递信息中不够精确的生理局限性,为人们提供精密的测量标准。人类的感覺器官往往带有模糊性,而且比较粗糙,不够准确,只能作定性的估计。借助观测仪器,可以把被测量对象的特征转化为我们视觉器官的感知对象,转化为测量仪器上客观的空间标度和数值显示,从而精确地测定对象的数量特征。随着科学向精确化和数字化的方向

发展,量的测量也越来越重要,对测量精度的要求也越来越高。通过不断改进仪器的精度,减少客体信息的衰减和失真,从而使它们获得更多,更精确的科学信息。观测仪器还可以克服人的感官在反应速度上的生理局限性,为人们提供准确的测量和记录手段。此外,观测仪器还可以帮助人们消除由感官造成的某些错觉,使观测的结果更为客观。

但是,观测仪器的认识功能也有一定的局限性。首先,通过观测仪器所完成的间接观测不如直接观测那样直观,必须把直接观测和间接观测有机地结合起来。其次,运用观测仪器也可能导致错误的观测结果,造成系统偏误。如果设计观测仪器的理论依据不充分、不严密,以及在制造和操作观测仪器时技术上的失误等,都可能导致错误的观测结果。

运用观测方法进行统计认识,是一个复杂的认识过程,观测的根本目的在于获得观测对象的感性认识材料。为进一步的统计认识提供源泉和动力,因此,观测材料的可靠性至关重要。为此,科学观测还必须坚持观测的客观性这一根本的原则。遵循观测的客观性原则,就是要使观测所获得的事实材料反映客观事物的本来面目。

美国科学哲学家汉森于1958年明确提出了“观测渗透理论”的重要命题。他指出:“由于各人具有不同的经验、认识和理论,具有不同的理论背景和知识,所以不同的人能从同一对象中观测到不同的东西;进行观测陈述必须运用某种语言,作出具有某种理性结构的判断。这样,就会导致两种可能:以正确反映客观事物本质的理论来指导观测,将保证观测的客观性;相反,当不完备的甚至错误的理论和观点渗透到观测之中,就会导致观测的主观性和易谬性。”(汉森:《自然科学哲学问题丛刊》,1981年第2、3期)。

观测的缺陷在于①感觉器官的生理局限性。但是,随着科技的发展,在观测工具和仪器的帮助下,感觉器官的间接性无限

地扩大了观测的可能性。②不能在严格对客体进行控制的条件下进行(不能控制某个过程的条件和进程,不能无限制地再现观测条件,不能在观测范围内改变客体。)。③统计认识主体的主观性。在观测中,由于有先入之见、无意过失、假象和错觉等因素,都会导致观测主观性。先入之见主要是统计认识主体错误地或过份地夸大某一理论、某一方法或自己的经验造成的。如果统计主体有意识地带着某种偏见对客体进行观测,就会严重影响统计观测所建构起来的信息系统与其客体系统的同构性。无意过失主要指观测者在观测过程中无意识地渗入了某种主观因素,使观测资料没有如实地反映客观情况,出现错误的信息描述。假象是事物本质所表现出来的一种现象,但却给人一种同事实真相不同的错误的感觉。往往和人们观测事物的角度有关。错觉的产生则和人的感官和心理因素有关。在统计观测过程中,一旦产生错觉,势必影响观测结果的客观性。

观测的这些或其他的不足之处在实验中将克服,历史上这种向实验的过渡是科学认识日益增长的需要和要求所引起的。向实验的过渡是在感性认识间接性的发展中新的步骤。

总之,进行统计观测,由于受各种复杂因素的影响,要使观测材料更加客观、可靠,并非简单。首先,要以正确的科学理论为指导,坚持实事求是的科学态度,防止和克服观测的主观性。其次,进行系统的全面的动态的观测,防止片面性。即从事物的空间分布上,观测事物的各个方面,事物的全体;从事物本身以及一事物与它事物的相互联系上,观测事物的整体特征和事物在环境中或更大系统中的表现;从事物的时间演化上,观测事物变化发展的各个阶段和事物发展的全过程。再次,要注意选择具有典型意义的观测对象和观测条件,避免次要因素的干扰。

### 3.2.3 统计实验方法

统计实验也是获取原始信息的基本方法,是形成和检验统

计认识成果的实践基础。统计实验也属于科学认识的经验层次。统计实验方法是统计主体根据统计认识的目的需要,运用一定的仪器和设备等手段,人为地变革、控制或模拟客体,在有利的条件下获取原始信息的一种研究方法。统计实验和统计观测虽然同是获取感性材料的方法,直接目的都在于建构客体的原始信息系统,但是统计实验与统计观测有着重要的区别:统计实验是实践的基本形式之一,是人变革外部世界的一种对象工具活动,而统计观测则具有直观性的特征。统计实验方法比单纯的统计观测方法具有明显的优点。单纯的统计观测一般只能在天然条件下认识观测对象,常常受到诸多条件限制,因而具有被动性的一面。而统计实验区别于单纯的统计观测就在于其本身的能动性质和对现象的自然进程的干涉。即统计实验是在有目的地变革客观条件中认识研究对象,便于取得在自然条件下难以获取的原始信息,大大加强了人们认识自然的主动性。此外,统计实验中必须使用统计观测方法,在某种意义上说,统计实验是一种有控制的统计观测。而统计观测方法除了在统计实验中运用之外,还可以在那些尚不可能运用统计实验方法来变革和控制研究对象的领域而得到广泛应用,使之比统计实验方法的应用领域更为广泛。

统计实验作为社会实践的一种特殊形式,有着同其他实践活动相类似却不尽相同的要素和结构。统计实验者、统计实验手段和统计实验对象是统计实验的三个基本要素。统计实验者作为认识主体是统计实验活动中首要的、活跃的因素。统计实验者的科学知识和统计实验技能的高低,对统计实验水平有着重大的影响。统计实验手段是把统计实验者的作用传递到统计实验对象中去并获得有关信息的各种统计实验仪器、工具和方法。它作为统计实验者和统计实验对象之间相互作用的“媒介物”,乃是人的感觉器官、劳动器官和思维器官的延长。

统计实验的认识功能主要表现为能够为人们提供重要的原



始信息。从而成为统计认识的重要基础和检验统计认识成果的价值的重要手段。

从统计认识过程来看,统计实验就是将统计认识客体控制在一定条件下,对统计客体进行重复(或大量)观察、测量,取得统计原始信息的方法。统计实验一般有两种情况,一是在相同条件下,对同一现象进行多次重复观测或试验,以取得认识对象样本数据的方法。二是为了研究各种现象之间的内在关系。通过控制影响因素,来取得被研究现象的有关信息,通常称之为有控制统计实验或直接统计实验;或者是通过观测与影响因素的不同水平相对应的被影响因素的取值,来取得统计原始信息,一般称之为观测统计实验或比较统计实验。例如通过观测不同收入水平组别的人的不同消费模式,来获得收入水平与消费模式之间依存关系的统计信息。

统计实验方法的优点表现为以下几个方面:①统计实验法是确立因果联系的最好方法。统计实验法引入自认为是原因的自变量的变异,观测其是否会导致因变量的变化及变化的具体表现。然而在某些社会现象中,能否运用统计实验法还有不同的看法;②统计实验法是一种高度受控统计实验,通过环境的控制,有利于撇开外部错误信息的影响,可以把被研究的客体同附属的、次要的、使其本质模糊不清的现象的影响隔离开来,并在纯粹的形势下进行研究;③在严格的控制下多次再现被研究的过程;④有计划地改变、组合各种条件,以便得到已欲探求的结果;⑤统计实验法是一种纵向的调查研究行为,能对调查对象在该环境下的发展变化进行跟踪记录。

统计实验法的缺点表现在:①某些领域,例如在对部分社会现象的调查分析中,统计实验法的结果与自然环境下的结果可能相去甚远;②统计实验法只能对目前现象进行统计实验,对过去和未来不可能或难以了解;③统计实验法受样本容量的限制,若样本过多,统计实验研究或控制外界变量的困难就越大;④统

计实验所需时间较长。且包含一定的风险,费用支出往往较大。尽管如此,统计实验法仍不失为一种卓有成效的搜集统计信息的方法。

### 3.2.4 统计测量方法

测量(Measurement),通常是指人们对客观事物进行某种数量化的测定,是一种定量的研究方法。其目的在于在人类认识中更完全、更精确地用数和量来反映客体的量的大小和关系。C.C. 斯季文斯指出:“测量从广义上来说是按确定的法则,赋予客体或事件以数的形式。而数的形式能根据不同的法则归属于客体这一事实导致运用测量的不同标度和不同类型。在它们的某一部分中,法则本身属于在进行实验时所实现的具体的经验步骤。”(转引自《科学认识的方法论问题》,知识出版社,第89页)。

统计测量是根据统计认识所指向的客体的特征的性质及可测性,按照一定的法则对客体特征进行赋值的认识过程。是获得客体数量信息的方法。从理论上讲世界上的事物都有程度上的不同,程度上的不同也就是数量上的差异。美国心理学家桑戴克(E.L. Thorndike)在1918年曾说:“凡物之存在必有其数量”(转引自陈选善:《教育测量》,河南大学出版社,1987年版,第2页)。美国测量学家麦柯尔(W.A. Mccall)于1922年又提出“凡有数量的东西都可以测量”(转引自陈选善:《教育测量》,河南大学出版社,1987年版,第3页)。这就是说,世界上的一切事物都具有可测性。但是,测量工具、测量方法、测量精度有所不同,它取决于客体特征的性质和人类认识能力及测量科学的发展。例如,就目前的认识水平和测量手段来说,离散性特征的可测性高于连续性特征的可测性;自然客体的可测性高于社会客体的可测性;而社会客体的可测性又高于思维客体的可测性;客体外延性特征的可测性又高于内涵性特征的可测性;物理特征

的可测性又高于观念特征的可测性。

统计测量是对客体特征量的规定性的标度。量是事物的规模、程度、速度;及其构成成分在空间上的排列组合等等可以用定性区别和定量确定(数量表示)的规定性。也是物质的一种存在形式。量与事物存在但并不直接同一,量总是存在着一定的变化幅度的。量的规定性是客观的、是事物所固有的。人们头脑中的数量概念不过是客观事物数量关系的反映。事物的量的规定性又可分为内涵的量和外延的量。外延量是表示事物存在范围和广度的量,是可以机械的方法来计算的量;内涵量是表示事物等级程度、构成方式、功能过程的量,它比外延量更加深刻,是不能用机械的方法来计量的。一般来说,内含量(如硬度等)以及主观状态、主观属性具有较低的可测性,而外延量具有较强的可测性。

内含性属性的可测量性一是指内含性的不同等级能够被排成序列,二是指借助可测量的外延属性,依据其相互之间的规律进行间接测量。

主观状态的特征,如评定知识、道德品质、才智、熟练程度、技巧、艺术技能、价值、方针等个体和社会团体的主观状态的特征至少到现在还没有创造出物理测量的工具。在为了获得这些信息所确定比较“标准”或“准则”时,本身就包含着主观性。如考试成绩的价值以主考人的学识水平、诚实和原则性、情绪、观点等这样一些偶然的情况为转移。

统计测量可以具体分为计数、测量、准测量、伪测量(量化)和计算等五种方法。

计数 是形成一般的数的最初方法,也是获得精确的定量信息的方法。它基于这样一个事实:由一个个元素组成的集合不变的属性就是个体的数目。计数的操作就是把集合中任意多的“一”合在一起。计数是以利用计数工具为前提,即以作为相互之间相对有序的大量元素的专门的数字和符号为前提。通过计

数得到的数值执行的是数字作为计数的工具的职能。计数的难点,也是关键,在于界定所研究客体。例如,计算一个单位或系统知识分子的数量,关键在于给出判断“知识分子”特征的客观标准。当有了这个标准,并予以区分之后,计数问题就迎刃而解了。可见,就连最简单的计量方法,也必须以定性分析为前提。

测量是以创立和使用作为测量手段的测量技术的一定的物理过程和物质工具为基础的活动,也是以某些理论前提为依据的活动。

测量——狭义测量的严格定义为“包括在用物理实验的方式将所给的量和当作标准的量(就它的某个方面的性质)加以比较的认识过程。表示被测量与单位关系的数称为被测量的数值,它可以是整数或者分数和抽象的数。当成单位的量的值称为单位的大小。”(转引自《科学认识的方法论问题》,知识出版社,第94页)

可见测量具有三大要素,即应该有某种用来同被测量的量作比较的客观标准、在物理上被证明是正确的实验程序、正确的量具。这些因素旨在尽可能最大限度地控制测量过程的随意因素和主观因素。

测量的形式化条件:

①给出  $n$  个物体  $b_1, b_2, \dots, b_n$  的集合。它们相对于某个属性的排列在下述条件之下是可能的,即有下列关系中的一种:  $(\alpha) b_i > b_j; (\beta) b_i < b_j; (\gamma) b_i = b_j$  (“ $<$ ”和“ $>$ ”象征着被研究的属性的客观差别)。关系“ $<$ ”是不对称的。

②如果  $b_i > b_j, b_j > b_k$ , 那么  $b_i > b_k$ 。即给定关系是可传递的。

③如果  $b_e + b_f = b_g$ , 那么  $b_f + b_e = b_g$ 。

④如果  $b_i = b_j$ , 那么  $b_i + b_j > b_i$ 。

⑤如果  $b_i = b_i', b_j = b_j'$ , 那么  $b_i + b_j = b_i' + b_j'$ 。

⑥  $(b_i + b_j) + b_k = b_i + (b_j + b_k)$ 。

同时满足上述 6 个条件,才可以对客体的指标进行严格意义上的测量。能否运用数学方法进行分析,取决于被测量的数字符合上述条件的程度。把被测量的数值引入数学方程式的时候,我们就能用它们进行所有的数学运算,我们将确信得到的结果反映出客观的量的关系。

在测量中,区分客体特征是否具有自然原点是非常重要的。对于任意原点来说,客体特征的测量值是相对于测量单位和假定原点的“相等的间隔”,因此,这里的数字执行着作为确定间隔的相等或差异的手段之功能。其数学结构可以表示为

$$x = a + M/b$$

其中  $x$  是任意数字,  $a$  为任意假定的原点,  $b$  是测量单位,  $M$  是被测客体特征的量。这里所作的测量操作是确认间隔相等。对于自然原点来说,即  $a = 0$ ,显然客体特征的测量值是相对于测量单位的“相等比率”,测量得到的数字具有表达量的关系的功能。相应的数学结构为

$$x = M/b$$

这时测量者所作的操作是确认比率相等。

严格地说,依据实验并满足上述形式化条件的量的评价可以被称之为测量(这些条件的满足也是实验证明了的)。在不能满足上述条件的情况下,不说是测量,而说是借助于标度、排序等来作量的和质的评价的各种手段和方式才是合适的。

准测量就是对所研究客体的大量个别事物或现象按照递增或递减规则进行赋值,以标明其属性差异的顺序、等级等。只有当客体满足测量的形式化条件的前两条规定时,才可以对客体进行准测量。准测量的数字的功能是作为标志在有序系统中某种性质的确定地位的符号。测量值标志着严格的位序差异,但并不表明事物之间的数量差异量。正是在这种意义上,我们称之为准测量。其数学结构可以表示为

$$x = f(E)$$

其中,  $x$  是任意数字,  $E$  是客体个别事物或现象的某一被测特征,  $f$  为单调递增或递减函数关系。其测量操作在于确认“<”、“=”、“>”。

伪测量即对所研究客体的大量个体或现象赋以对应的数字, 以标识不同的事物。在赋值时, 既可以对每个个体不论其性质是否相同, 均赋以一个不同的数字, 这实质上是对事物进行编号; 也可以对性质相同的个体或现象赋以相同的数字, 则此时数字具有标识不同类别、不同组别的功能。显然, 不论是前者, 还是后者, 这里数的形式所完成的只是命名的功能, 而不是计算的功能。——也就是说只有数的符号之名, 而无数的大小之实。目的仅仅在于标识。就是对于后者来说, 我们也只能对其所包含的相同事物进行计数, 而不能进行任何数学运算。这里的统计数字只具有数字作为命名的手段、特殊的标志或恰当的鉴定同质的记号的功能。这里遗漏号码是可以的, 因为号码的顺序不起重要作用, 被编号的客体之间的关系既不符合被归属于它们的数字之间的顺序, 也不符合被归属于它们的数字之间的量的关系(序与量)。根据广义的测量是“按确定的法则, 赋予客体或事件以数的形式”这一点来看, 上述量化方法无疑具有测量的意义。但是, 这里“赋予数”的时候具有很大的随意性和主观性, 并且“数”又不具有大小和多少的性质, 因此, 我们称之为伪测量。如果要建立其数学结构, 则有

$$x = f(T)$$

其中,  $x$  是任意数字,  $T$  是客体的个别事物或现象,  $f$  表示一一对应关系。有了这一数学结构, 我们可以根据认识的需要对客体的测量结果进行方便的替换, 即

$$y = g(x)$$

这里  $y$  是任意数字,  $g$  表示一一对应的函数关系。伪测量的最基本的测量操作就是“确认相同”或“区别不同”。

计算就是根据已知(测量)的量及其与未知量之间客观存在

的数量关系,通过相应的计算操作,对客体的被测特征进行赋值。也可以称之为间接测量。对于一些内含性属性不能直接测量,但可以通过把内含性属性同可以直接测量的外延性属性联系起来的定量规律而实现。如密度是一个内含性属性,若直接测量只能得到不同物质密度的一个序列,因为数的归属是任意的,而序列的关系仅仅满足不可逆性和传递性的条件,但是可以根据密度和质量、体积的规律,通过质量、体积间接地测量密度,这时赋予密度的不同等级的数不是任意的,而是由质量和体积之间的客观联系得到的。在这个意义上说,密度是可测的。这种间接可测性取决于内含性属性与外延性属性是否客观存在一种能够联系在一起的量的规律。

此外,对于一些复杂客体的外延指标有时也不能直接测量,如社会产品总量、国民生产总值等等,需要借助其它可测量以及它们之间存在的内在的数量规律,经过计算进行间接测量。

综上所述,不同的测量得到的测量数字从形式上看是一样的,但是其所执行的数字功能是完全不同的。因而对统计测量得到的原始数字施以统计计算处理时,所能运用的数学运算方法是不同的。统计主体在对原始统计数字进行统计计算时,必须考虑原始数字的测量方法,根据其所具有的数字功能选取相应的数学运算法则进行计算操作。否则,所进行的数学运算及其结果是毫无意义的。这也是统计计算区别于数学运算的重要之点。

统计测量的客观性是相对的。没有绝对“客观的量化测量”。因为没有主体的干预是不可能的。这主要表现在:①在进行统计测量之前,统计主体首先要区分出被测量。这个量需要澄清偶然影响、障碍、歪曲。例如,在测量土地面积时,可以避开沟、坑、梁等的不平整性,将其想象为象几何平面一样平整。②自然要受到测量工具(测量工具是研究者对被研究客体能动的和无限扩大的干涉的必要手段)、方法的影响。③测量过程往往

具有一定的近似性。④建立测量标度方法也存在着一一定的近似性。

### 3.3 统计信息加工方法

#### 3.3.1 统计描述方法

统计认识不能只是直观地描述原始信息,门捷列夫说:“单是事实的收集,哪怕收集得非常广泛;单是事实的积累,哪怕积累毫无遗漏,都不能使你获得掌握科学的方法,不能向你提供进一步成功的保证,甚至还不能使你有权按照科学这个名词的高级意义来把它叫做科学。”(转引自费而斯曼:《门德列也夫周期律在现代科学中的应用》,中国青年出版社,1959年版,第38页)统计认识不能停留在经验层次上,必须对原始信息进行理论思维的整理和加工,使科学认识发生质的飞跃,深入到理论层次,才能揭示客观现象的本质和规律。爱因斯坦曾经指出,“知识不能单从经验中得出,而只能从理智的发明同观测到的事实两者比较中得到。”

统计主体在得到客体的原始信息之后,就要将统计方法施行于原始信息,进行一系列的信息转换,以获取客体基本特征的有关信息。加工原始信息的方法,与获取原始信息的方法不同。观测、实验等获取原始信息的方法,是在人们的感性活动中实现的,所以通常称为感性的方法或实践性方法。加工信息的统计方法,则是根据数理逻辑和统计法则,并结合统计主体对客体的现有认识通过理论思维活动和统计计算进行的,所以通常称为理论的方法或理论性方法。

##### 一、统计分组

统计分组是根据研究的目的和客体的特点,按照某一分组标志将总体分为若干个性质相异的组成部分的一种统计认识方



法。它是哲学上的比较与分类方法结合运用的具体形式,可以反映客体在分组标志下的差异结构及其分布状态和分布特征。同时也是使零散的资料系统化的一种具体方法。统计分组的原則是组内相同、组间相异。因此,它是一个“分”与“合”的过程,是根据所建立起来的信息系统对客体进行的一次定性认识。相对于总体而言,它表现为一个“分”的过程,是“求异”,将总体分解成若干个性相异的组成部分,每一个组都是一个低一级的小总体;对于各个组来说,它又是“合”,是在“求同”,将性质相同的个体结合在一起。

统计分组理论的内容十分丰富,但其中心问题是选择分组标志、划分各组界限和确定分组体系。选择分组标志是统计分组的核心问题。分组标志一经确定,必将突出总体在该标志下的性质差异而将总体在其他标志下的性质差异掩盖起来了。如,将某地区人口按性别分组,则只能反映该地区人口在性别方面的差异、结构及其分布;而年龄、民族、文化程度等方面的差异全都掩盖起来了。因此,分组标志的选择,不但直接影响统计分组和统计整理的科学性,而且最终影响统计认识的客观性、全面性和深刻性。在进行统计分组时必须根据认识目的,结合具体条件,选择其中最重要、最能反应现象本质的标志作为分组依据(标志),这是正确选择分组标志应当遵循的基本原则。划分各组界限就是要在分组标志的变异范围内划定各相邻组间的性质界限和数量界限。总体经过分组后得到的各个组,实质上是一些性质不同的小总体。从这个意义上来看,组限是质量互变的转折点,组距就是哲学上的“度”。度是物质和量的统一,是事物保持自己的量的限度、幅度、范围,是和事物的质相统一的数量界限。在度中,质和量处于不可分离的统一中。首先,度是质和量的互相结合。一方面量中有质,度中的量不是单纯的量,而是具有一定质的量;另一方面质中有量,度中的质也不是单纯的质,而是具有一定量的质。其次,度又是质和量的相互规定。质

和量的相互结合和相互规定,使质量双方在特定的度的范围内处于统一状态,形成某物之所以为某物的质和量的统一体;一旦某物的质和量的统一体发生分裂,也就是度的超出或破坏,某物就会转化为他物而形成新的质和量的统一体。

如果组限划分不当,必将混淆各组的性质差异。因此,在划分组限时,首先要研究总体在分组标志下有多少种性质不同的组成部分,然后再深入研究各组之间的界限。统计分组体系是根据统计认识的要求,对同一总体按照多个分组标志进行多种不同分组而形成的一种相互联系、相互补充,能从不同角度加深对客体数量表现的认识体系。是决定统计认识全面性的最重要的环节。

## 二、平均法

平均方法是从个别到一般的方法。它优越于简单枚举法和外推法。平均法不需要强调每个个体完全一致,才能推出一般。客观中个体之间差异的存在是绝对的、经常的,统计方法不是对此束手无策,而正是利用这些差异推断一般。在现象的合成原因中,结果既受到规律性的影响,也受到偶然因素的影响。规律性是不变的原因,偶然因素是变化不定的原因,事物的结果表现为这两类原因的结合。且往往固定原因总是被偶然原因所掩盖,用普遍的分别观测方法就不能确定固定原因的规律性。因此,每一个个体、每一次观测的具体结果不会完全相同,在此简单枚举法黯然失色。平均数法可以消除偶然因素的影响,显现出内在的必然性。

平均数法的归纳法表现为:随着试验(观测)次数的增多,平均数受到每一个观测值变化的影响,然而却不会明显地改变平均数。穆勒探索了因果联系的四种方法:求同法、求异法、剩余法、共变法及求同求异共同法(最后一种不属于基本方法,实际上是一种间接求异法。)。统计学的平均思想就是求同法“本质表或存在表”的具体应用。求同法应注意多选取有差别性的场

合,如果场合数目太少,会出现一些偶然的共同情况。适合于通过大量观测法去探索共性和规律性。

统计学的差异分析就是差异法的“差异表或接近中心缺乏表”的具体应用。求异法适合于科学试验的控制观测法。

### 三、相关分析和回归分析

从推理规则上来看,相关分析是共变法的具体化。共变法是说:如果某一现象发生一定程度的变化,被研究现象随之也发生一定程度的变化,则前者可被断定为被研究现象的原因。根据这一推理规则,在统计认识中,通过观测两个客体的数量变化关系,来判定客体间是否存在一定的依存关系。统计相关系数公式具体地实现了这一推理过程。

$$r = \frac{\sum (x - \bar{x})(y - \bar{y})}{\sqrt{\sum (x - \bar{x})^2 \sum (y - \bar{y})^2}}$$

每一次观测,是一种场合, $N$ 次观测就是 $N$ 个场合。观测值 $X_1, X_2, \dots, X_N$ 就是某一客体的 $N$ 个有关情况, $Y_1, Y_2, \dots, Y_N$ 是另一客体相应的 $N$ 个变化。如果 $X$ 的变化导致 $Y$ 发生相应的变化(统计上的显著变化),则认为 $X$ 与 $Y$ 具有相关关系,反之,根据现有观测信息无法断定二者存在相关关系。

回归分析是“量化”的求同求异共用法。在进行回归分析时,首先运用求同法求得回归方程,然后再用求异法验证:当解释变量不存在时,回归方程也不存在。从而确认是否接受所拟合的回归方程。

残差分析的原理是剩余法。穆勒的剩余法就是已知被研究的某一复杂现象是由另一复杂现象所引起,把两者间确认有因果联系的部分减去,则其余部分也必互为因果。在统计分析中,将这一规则进一步发展:如果影响 $Y$ 的因素只有 $X_1, X_2, \dots, X_k$ ,那么,从 $Y$ 中剔除这 $k$ 个因素的影响后所得到的剩余 $e_i = (Y_i - \hat{Y}_i)$ 就应该与 $Y$ 无关。所以,残差分析就是检验 $e_i$ 与 $\hat{Y}_i$

是否独立,若独立,则表明回归模型有效地刻画了各影响因素对 $Y$ 的影响,剩余是随机的;若不独立,则说明剩余 $e_i$ 不是随机的,可能会有影响 $Y$ 的其它因素存在。

#### 四、统计模型

模型是借助实体、图形或符号等手段,以某种程度的类似再现客体系统(原型)的系统。并在认识过程中以它代替那个原型,以通过模型研究得到关于原型的信息。原型(Proptotype)就是客观存在的具体事物、实际的研究对象、初始系统形式等,也就是统计认识客体。模型是对所研究的客体系统的结构或过程的一种简化、抽象和类比的表达方式。在认识过程中,往往一系列客观条件和认识客体的特点迫使必须用模型实验代替真真实实的研究客体。如客体在时、空上非常遥远、范围很大、不能直接实验,直接实验代价太大、对象过分复杂等。在所有这些情况下,借助模型实验是适当的,模型可在某种精确度上代替和再现真实客体,运用模型作为研究的专门手段的认识活动称为模型研究。模型一般分为物理模型和思考模型两大类。物理模型是以客体系统的结构和构造作为模型的组成元素,据此用缩小了(或放大)的尺寸,制作与实物系统相似的模型。模型的变量与客体系统的变量完全一样。如各种试验模型,都是物理模型。思考模型则是在认识客体系统之后,根据一定的逻辑变换规则而建立起来的一种刻画系统结构、特征及运动过程的一种表达方式。

统计模型是一种思考模型。它是根据一定的观测、实验(试验)或测量得到的原始信息,以及理论见解和可能的假说,运用统计方法和统计思维规则创造出的再现客体(原型)的某些关系或联系、状态或过程的准客体,哲学上也称设计物。它是客体的本质、结构、运行相依性以及发展变化规律等特征的统计描述。这种描述只能是真实系统的简化模型,而不是真实系统的本身。我们不能对准客体直接进行观测和研究,但可以在逻辑数学推

论的基础上和它们接触。用统计语言表述,统计模型中的模型就是观测值所属的随机变量的分布函数。一旦确立了观测值的分布函数,就能重新把观测值计算出来,这就是“说明已知的经验数据”的过程。当然计算结果与实际观测值一般会有误差,我们追求的是误差足够小。分布函数与观测值的关系,象似科学理论与经验事实的关系。科学理论是否被接受,要看它说明经验事实是否成功,成功到什么程度,正如统计模型是否成功,要看它与观测数据是否拟合,拟合到什么程度;为观测值建立起模型,人们不仅能描述对象的状态,还能进而预测未来的观测值,就象科学理论不仅能说明已知事实,也能预言未知的事实一样。

统计模型按其量化层次不同可分为两大类:一类是统计逻辑模型。这类模型是按照系统的逻辑结构,用框架图表的形式来反映和刻画所研究系统的特征、构成要素及其相互间的关系。如各种统计指标体系、统计分组体系等等。另一类是统计数学模型。这类模型是按照现象之间客观存在的数量依存关系,用一个或一组数学表达式来反映和刻画现象间的数量关系。统计数学模型一般分为回归模型、计量模型、投入产出模型和时间序列模型四种。回归模型表示被研究现象与其主要影响因素之间的数量关系。计量模型是用一组联立方程模型来研究变量之间错综复杂的数学关系。投入产出模型是用数学方程式的形式来反映客观经济过程和经济结构的一种方法,可以反映国民经济各部门生产的产品和提供的劳务与各项使用间的平衡关系,以及国民经济各部门在进行经济活动时,各种消耗与总投入间的关系。时间序列模型具体描述系统的动态结构和发展变化规律。

按照不同的分类标志,统计数学模型还有各种各样的分类。

按照模型的目的来分,统计数学模型可分为功能模型——用于研究系统的功能;结构模型——表明系统的构造;评价模型——评价系统的状态和功能的可信性和有效性;预测模型——

用于预测系统未来时刻的行为。

按照各变量所属时间是否一致来分,统计数学模型可分为静态模型和动态模型。当被研究现象和诸影响因素固定在某一时间条件时,被研究现象在某一时刻上的行为受同一时刻各影响因素的影响,没有时间因素的作用,这种模型就叫做静态模型。当模型中含有时间因素变量或者包含不同时间条件的变量时,模型的输入与输出表现为时间的函数,称这种模型为动态模型。静态模型可以揭示特定时间条件下,现象之间的静态依存关系。动态模型可以描述现象的发展变化规律性以及影响因素之间关系的变动规律性。

按照变量的性质分,统计数学模型可分为确定模型和随机模型。即各变量的变化是确定的并且它们之间的依存关系也是确定的,也就是没有任何随机因素的加入,这种模型称为确定模型。当现象的影响因素是随机的或者还受某些未知的因素随机决定,所研究现象与影响因素之间的关系不是一一对应的函数关系,表现为一种不确定的依存关系,这种模型称为随机模型。

按照现象与其影响因素之间关系的性质分,统计数学模型可分为线性模型和非线性模型。当模型表现为一次时,称之为线性模型;否则即为非线性模型。

统计模型与物理模型不同。物理模型是依照实物的形状和结构按比例制成的实体;而统计模型既不表现为实体,也不是所研究客体的一定比例的产物,而是一种比较完整、客观的近似刻画,既是人们认识事物的手段,又是人们对事物认识结果的描述。因此,统计模型与客体系统的符合程度取决于人们的认识能力、认识程度以及拥有客体信息的数量和质量。随着人们认识能力的不断提高和认识程度的深化以及新信息的不断获得,统计模型逐渐向其所描述的客体系统逼近。其次,物理模型与原型的关系是客观地、不依赖于人的意识而独立存在着,主体的意识和自觉性仅限于选择适当的模型,获知相似的条件,并在创

立或选择模型时使用这种知识。当这样的模型成为研究的客体之后,它就象任何一个物质客体一样,按照自然界的客观规律发挥作用。因而它们可以充当科学实验的手段,作为对象——工具的物质的活动形式。统计模型不仅是作为认识的固化手段履行一定的认识功能,而且是作为主体对客体的理解,作为想象的、理想化结构的认识功能。这些结构先是统计主体根据自己的认知结构和对客体的映象,在头脑中创生出来的,并且对这种结构进行着思想操作和信息操作。一方面统计主体的思想变革和思想的“实验”模型与原型同构、同态(Similar state,即模型与原型相似)、类质同形;另一方面客体的信息影响、决定着模型的拟合程度,并提示进一步改进模型的方向,启发、修正、引导着主体对客体认识的思想变革。可见,统计模型是客体信息与主体思想的总和,是理论与现实之间必要的中间环节,现实方面近似地反映在相应的模型之中。此外,统计模型的精度还受到随机因素的影响。统计模型作为一种认识手段,应尽可能地消除随机因素的影响,增强模型解释系统的能力;作为一种认识结果,统计模型包括已经认识和尚未认识的两部分,或者说是系统的真实关系与随机影响两部分。

统计模型具有复制性、相似性、代替性和外推性等特点。复制性——统计模型是建立在一定的统计认识客体系统的基础上而制成的,是原型的复制品。严格地说,不是完全拷贝,而是基本结构或某些特征的构建。使模型与客体(原型)保持同一是模型的客观性的基础和根源,也是其它特性的前提。否则,将会改变、歪曲模型与客体的关系。相似性——由于统计模型是客体的复制,因此,一方面统计模型的形式被客体系统客观地决定了;另一方面统计模型与其相应的客体之间保持着一定的相似性。这种相似既可能是外型相似,也可能是机理相似、结构相似、功能相似等,明确地描述和刻画了客体系统的结构、机理、功能。这使得模型与语言系统区别开来了。代替性——由于模型

具有复制性和相似性,因而使得模型在认识上具备了代替原型的能力,在统计认识过程中,它是原型的代替者。这一特点把统计模型同其它认识方法、假说区别开来了。外推性——根据建立起来的统计模型可以获得客体的有关信息,同时,在一定条件下,利用统计模型可以根据已知推断未知、根据过去和现在推知未来。

模型研究的直接对象是模型,而研究的结果又被转移到原型上,因此,对于这种转移的正确性的理论、论证乃是这种认识研究的必要条件和组成部分。

在物理模拟领域中,模型实验的理论基础是相似理论。相似是事物之间近似相同的现象。是一个现象范畴。相似有两个方面的含义。一方面,由于宇宙是由物质构成的,因而尽管宇宙之间的事物千差万别、千变万化,但总可以从中发现大量的相似现象。物质越是微观,其结构越是相近;若表现为宏观,就是其物理、化学性质越相似。另一方面,某种事物在发展过程中必然存在同质和变异。只有同质,才能保持同种属;只有变异,才能有所发展。相似是发展着的事物同质与变异的统一。所以,相似既存在于不同种事物之中,也存在于同种事物之中。相似与同一是两个不同的范畴,前者属于现象范畴,后者属于本质范畴。统计认识寄希望对客体本质的同一能够客观地把握,但是,人们能够直接得到的只是现象的相似。那么,究竟相似与同一之间有无关系?有什么样的关系?怎样才能得到这种关系?这是统计认识主体最关心的问题,也是统计认识结论是否客观、可靠的关键问题之一。相似原理告诉我们:相近事物都是由相似的单元、层次排列组合而来的;相似的基因、相似的条件和相似的环境产生相似的结果;事物包含的与相关事物的相似功能越多,其作用就越大,应用就越广。相似三定理中,第一定理表明相似系统应有相同的判据(系统的有关变量和参数之间所形成的无量纲比值),其相似指标应等于1。第二定理假设任一物理



系统是由  $n$  个量纲不同的物理量所组成的函数关系来表示,而且其中  $k$  个是相互独立的,则表示这一物理量的方程式可以用  $n-k$  个无量纲的比例数(即相似判据  $\pi_1, \pi_2, \dots, \pi_{n-k}$ )完全表达出来:

$$f(\pi_1, \pi_2, \dots, \pi_{n-k}) = 0。$$

第三定理规定了现象相似的充分条件:如果现象的单值条件(系统的几何性质、介质的物理性质、起始条件边界条件等)相似,并且由这些条件所组成的相似判据在数值上相等,则这些现象是相似的。三个定理作为一个整体,要求在进行物理模拟实验时,应该特别注意测量相似判据或与相似判据有关方面的量,如果测到了这些数据,就可以根据相似判据之间的函数关系式来整理它们。当要把认识结果应用到其它同类现象上去之前,还必须对各现象的单值条件是否相似、相似判据是否相等进行检查,如果满足要求,则结果可以类推。

同质相似性:同类事物之间的相似性。所谓同类是一个相对概念,是相对于划分方法而言的。根据系统层次结构的观点,系统的内部结构与外部结构随着客体系统外延的变化而相互转化。在一个大范围内属于同一类的事物,在一个小范围内就可能不属于同一类;反之,在一个小范围内不属于同一类的事物,在一个大范围内则可能属于同一类;任意两个事物,只要存在共性就有相似性可言,相似程度可以通过相对距离去界定。设  $A$  到  $B$  关于范畴  $p$  的相对距离为:

$$D(A, B)_p = \frac{F(B) - F(A \cap B)}{F(B)}$$

其中  $F(B)$  为  $B$  的特征的数目,  $F(A \cap B)$  为  $A$  与  $B$  的共同特征数目,则  $A$  对于  $B$  关于  $p$  的相似度即为:

$$R(A, B)_p = 1 - D(A, B)_p$$

相似理论仅限于同建立同质现象之间和属于物质运动的同一种形式的系统之间的关系。

在数学模拟中,模型与原型关联的基础是推广了的相似理论。它不考虑模型与客体在质上的差异性和它们对于物质运动的不同形式的隶属性,这种推广采取了系统类质同形的更加抽象的理论形式。

类质同形是这样定义的:有两个集合  $D_1$  和  $D_2$ ,只要满足以下条件,则它们相对于元素  $a_1 \cdots a_n$  和  $b_1 \cdots b_n$  以及关系  $R$  和  $S$  是类质同形的( $R$  和  $S$  是按照  $D_1$  和  $D_2$  定义的):①存在着一个映象函数  $F$ ,在它的作用下,集合  $D_1$  中的一个元素对应于集合  $D_2$  中的一个且仅仅是一个元素,反之亦然。②此外,不论什么时候,只要集合  $D_1$  的元素处在关系  $R$  中,那么它们的  $F$ ——映象将处在关系  $S$  中,反之亦然。如果它们在所有关系中类质同形,那么这两个集合就完全类质同形(Apostel L. Towards the Formal Study of models in the Non-formal Sciences.——The Concept and the Role of the Model in Mathematics and Natural and Social sciences Dordrecht, 1961, P17)

在数学模拟中,模型与客体之间有这样一种关系,这种关系使可能把对模型的实验研究结果归属于原型的相应属性。例如,为了再现机械系统的相应属性而建立电模型,乃是一种结构层次上的类比,它建立在不同的自然规律有同一数学形式的基础上。因此,以系统的类质同形为依据的物理类比方法,是相似理论对于本质上是属于物质运动的不同形式的系统之间的关系的情况的推广。如果在相似理论中不变式或判据、模型和样本的相似,是根据同样的自然规律(物理规律)的作用推导出来的,那么根据不同的物理规律具有同一数学形式这一事实,类比学说或物理过程的类质同形理论将导出相似标准。这里研究的已不仅是一种运动形式范围内的不同系统,而且研究的是作用于不同自然领域中的不同的物理规律。

在彼此相似的系统中:①包括在一个系统中的元素与包括在另一个系统中的元素,它们质上有别,但彼此单值地对应。

②一个系统中原素之间的关系与另一个系统中元素之间的关系彼此单值对应。这种对应清晰地表现在数学形式的同一性中。

方程式的数学形式是对这些普遍规律的反映,是它们独特的符号模型。由于这些客观存在的运动规律的同一性,在对规律进行相应解释的条件下,这种模型对于所有具体的运动形式是类质同形的。

可见,一定类型运动的普遍规律的存在是模拟的客观基础,也就是进行模型实验和把模型实验的结果转移到同模型相比具有另一种物理本质的相应客体的客观基础。把相似理论推广到物理类比(在类质同形和同态概念中被逻辑地明确化了)的理论,是在这个范围内模拟的理论基础。

在控制论中,模拟的方法比在数学中的模拟具有更普遍的性质。一方面控制论中模拟方法的应用范围扩大到生物学和社会学的系统和过程的领域中,另一方面它不仅撇开了组成控制论系统的元素的不同本质,而且还撇开了这些元素彼此用以相互联结的具体方式,因而撇开了这些系统的特殊规律,控制论把注意力集中在控制和自组织系统发挥职能的普遍规律上,而不管这些系统是由人工建立的技术装置,还是生命机体或人类社会。

统计模型是统计认识的进一步深化,是统计思想、统计理论和统计方法的飞跃性的发展。其理论基础是信息与客体的同一和数学模拟中推广了的相似理论。前者是统计模型能够刻画客体特征的本质决定,后者保证了在建立统计模型的过程中数理推理的同质不变性。电子计算机的普及和发展,为统计模型应用开辟了广阔的领域。统计模型是贯穿统计认识全过程的基本方法,是统计分析的最普遍、最严密的方法。

统计模型的主要作用可以概括为三个方面。一是描述现象之间的统计关系。这是统计模型的基本作用,不论是统计逻辑模型,还是统计数学模型都具有这一作用。二是对某一系统的

行为或运动过程进行评价。三是统计模型可以用于预测和控制。统计模型在静态上是对所研究现象与其影响因素之间内在依存关系的客观描述,在动态上是对现象以及想象与现象之间的发展变动规律的刻画,预测就是这种统计规律性的延伸和发展。根据统计模型所揭示的统计关系,一方面通过对主要影响因素的调节和控制来校正系统的行为,另一方面通过修改参数,可以调控系统的运行过程和结果即控制。

统计模型方法的生命线在于建模的理论依据(相似理论、类质同型)的正确性和模型研究结果对原型的性质的符合程度。值得注意的是解释结构的模型与描述行为(或功能)的模型其形式不是时时保持一致的。因此在应用模型时,至少在两种情况下被认为是成功的①如果模型展现出与原型的行为相似的行为,②如果在对这个模型的行为和结构进行研究的基础上,能够揭示出新的关于原型的特征。

### 五、统计概率

在本体论中,概率是客体的客观可能性的量度,是客观世界的一种现实属性。表征的是客体的客观特征和规律性;这种属性是大量现象的特点。规律性表现为透过偶然的振动、偏差、涨落而达到占优势的趋向。实验、试验和观测是确定概率的手段。因此有时也称统计概率、经验概率。当概率表征单个微观客体的某一行为在发展变化中的可能性时,概率表现为过程的性质,是客体某一行为在一定条件下出现(发生)的现实的、客观的可能性,具体体现为已实现的情况的频率。可见,概率表述大量随机现象的客观属性或者在一定的外部条件下,单个客体行为的不确定性。

逻辑的概率则表现为用前提“证明”“结论”的“程度”的可能性,表征的是陈述之间的逻辑关系。例如:有两个陈述:“在这片树林中百分之九十的树是松树”(A)和“在这片树林中这棵树被砍掉了”(B)。由(A)和(B)得出第三个陈述,被砍掉的这颗树是

松树(C)。”这时,用前提(A)和(B)证明假说(C)的程度的逻辑概率为90%。作这样的表述显然从逻辑必然性的本质那里找不到理论依据。我们必须透过逻辑结构,研究反映在逻辑结构中的客观联系。如果说逻辑必然性是脱离现实之偶然的、个别非本质方面的客观必然性和普遍性的抽象,那么逻辑概率则总是同大量这样的偶然性和个别现象联系着,必然性就是在大量偶然性和个别性的形式中实现的。可见,反映陈述之间逻辑关系的逻辑概率是以陈述含义为基础的,而不是建立在实际情况基础之上的。在归纳逻辑中,前提证实结论的程度,取决于对相应陈述含义的逻辑分析,而不是取决于观测、实验、测量等经验程序(方法)。经验程序(方法)的目的是确定相应的事实。这里需要注意的是认识论的分析与逻辑分析不同,前者在揭示陈述是否符合观测和实验中所确定的事实,即它们是真实的还是虚假的,而后者则只限于确定它们的含义和考察陈述之间的联系,这种联系建立在它们的理性含义的基础之上,而不是建立在实际内容和客观真理性的基础上。正如卡尔纳普曾说的:“只要求我们能够理解以这些含义为基础的它们的某些关系。”由于陈述反映着我们关于客观事实的知识,因而逻辑概率无法排除和代替我们对陈述进行认识论的分析。

概率的经典定义是建立在所有事件的绝对对称和具有相同可能性的基础之上的,概率值是有利于预期事件的随机数与所有等概率的且(和)互不相容的随机数的关系。这是拉普拉斯决定论的表述。这只有在理想模型的情况下才能满足。概率的统计解释更加适合反映大量现实事件的规律性。它是把作为一门纯粹数学学科的概率论的公理体系同客观实在性、同在我们周围发生的、以概率规律性为特征的现实过程联系起来的环节。

## 六、结构分析

由于系统的功能首先取决于系统的结构,因此,统计认识的重要任务是弄清客体系统的有序结构,人类历史上一一直在探索

的就是一切客体系统的结构。结构研究方法属于形式化方法的范围(即使形式暂时摆脱内容,以系统形式为特殊对象的形式化研究方法),它的基本任务是研究系统结构的有序形式,从量的角度去研究系统。具体说来,有以下主要任务:

1. 探索客体的有序结构。统计分析客体结构有两种途径:一是功能分析,一是解剖分析。在结构未能弄清之前,我们只能知道系统的一些信息、能量、物质的输入值和输出值,而不能确切地知道它的内部结构及内部变化。这时,只能通过系统的功能分析去判断系统的结构。事实上就是黑箱方法。如古代中医学的“望闻问切”方法就是先获得病人输出的信息,然后医生经过对病人各种信息的综合分析,即根据病人功能所发生的种种变化,从而进一步判断病人身体内部结构所发生的病变。这时,人体结构对医生来说就是一个黑箱。

当已经掌握了系统结构机理,并且可以得到有关元素及其相互作用的统计资料时,才可以进行结构复合、结构还原。

2. 测定有序度。非平衡是有序之源。系统通过涨落,从环境中吸入负熵流,而使系统从原来的无序状态走向一种时间、空间或功能有序的动态结构—耗散结构。

3. 测定系统从环境吸纳的熵

4. 结构分析的另一重要任务就是用一种结构去模拟另一种结构系统的功能即功能模拟方法。

5. 改变系统结构以提示系统的功能。

6. 选择最优结构。

### 3.3.2 统计推断方法

统计推断是统计主体根据样本构成的信息作出关于总体构成结论的一种推理过程。推理是以判断为中介并推出判断的过程,是判断体系。这个体系由三类知识构成:基本知识(包含在推理前提中的知识),推出的知识(通过推理过程获得的知识),

论证的知识(确定可以从前提过渡到结论的知识)。表现为论证的知识有公理、规则、定义、规律、以及其它一些具有确实性或概率性的原理。论证的知识决定着推理的形式,决定着从前提过渡到结论的性质。这种知识对于包含在前提和结论中的知识总是一般的。因此,推理过程总是通过一般并在一般的基础上,即在现象的合乎规律性联系的知识基础上进行。所以,自然和社会中存在的客观规律性就成为推理过程即由已知向未知过渡可能性的基础。从个别过渡到一般和从一般过渡到个别,以及从具有一定程度共性的知识过渡到具有同等程度共性的知识都是在这规律知识的基础上实现的。

推理形式具有相对性。推理形式的正确性不依赖于前提的真实性,有其积极的方面,它使推理具有能动的创造性。为了获得真理,不仅要从真实前提,而且要从虚假前提中作出形式上正确的推理。如果推理只能从那些真实性早已知道的前提出发,那么推理就不成其为获得真理和证明真理的能动力量。

统计推断就是根据已有的既得知识、业已观测到的现象、或进行过的实验,运用统计方法进行推理,获得新的发现。

统计推断是归纳推理的特殊形式,即由已知的统计事实到未知的概率结论的形式。统计推理是运用统计方法于归纳推理之后而得到的完善化的归纳推理。它的本质是,为了得到关于总体属性的完全信息,并不必研究它的所有个体,而只须研究它们的一部分,即样本。问题在于如何提高(保证)推断的可靠性。显然简单地回答就是要建立样本与总体之间相似性的本质联系,使样本尽可能地再现和表征出总体的本质属性。具体说来,一般要满足以下三个要求(这是由于统计归纳推论中,被研究客体的非单质性,即属性概率的“离散性”决定的):①增加样本单位数目,②代表性的要求,意味着样本应当最恰当地再现总体的结构,这是外推可靠性的最重要的依据之一。这个要求使得可能把样本看作总体的模型。由于在结构方面的相似(也就是类

质同形),或在其他被分出的关系方面总体相似,对样本的研究将提供得到总体信息的可能性。因而,在这里认识的情况同模拟方法中一样。③抽样的要求,以保证样本(统计模型)的客观性。抽样推断实质也是一种模型研究,样本即为模型,抽样即为构造模型(由原型到模型),推断为由模型到原型。因此,可用模型的相似性和可推断性来研究或丰富解释。这里,就增加样本容量的本质和作用展开深入探析和研究。

### 一、增加样本容量的哲学意义

增加样本容量旨在增加差异性。我们说样本来自同一总体,要求构成样本的每一个样本单位都是所研究总体的一个单位,也即具有同质性,现在的问题是根据对一部分(样本)单位特征的观测研究,以一定的合理信度得出该总体也具有这种特征的结论。显然,为使所得到的结论更加可靠,希望所建构的样本尽可能与总体同构,通俗点说,就是要求样本除了“都属于同一总体”这一根本的同质性以外,其它方面的属性尽可能地保证其差异。(凯恩斯称之为“负相似”),这种差异越大,所得出的相似性之间的“本质联系”越可靠。相反,差异越小,结论越不牢固。例如,我们研究公众对某一事件的态度。样本单位除了是“公众”这一共同属性以外,在年龄、受教育程度、职业、民族、宗教、性别、生活区域等方面尽可能不同,这样得到的样本起码在构成要素(零部件)上与总体接近了一致。得到的统计关系才能在总体范围上成立。否则,得到的统计关系仅适合于该样本,而不一定适合于总体。假设一种极端情形,在上述问题中,我们所调查(构建的样本)的都是大学生,那么由此得出的态度,仅属于“大学生”,而不属于“公众”。进一步假定构成样本的是“某一相同专业的大学生”,那么所得到的结论(态度)仅属于该专业的大学生的态度,而对于所有大学生来说,也不一定能够成立。还有更极端的情形,假设这个样本是通过总体的某一个单位多次重复抽取得到的(有放回取样),那么这个结论(态度),仅是该学生个



人的态度,而丝毫不能代表该专业的大学生的态度,更不能代表大学生、公众的态度。问题发生在哪里呢?出在样本单位缺乏差异性。可见,样本单位的差异性是提高统计推断结论信度的重要途径。在统计推断中,常常强调增加样本容量(增加实验次数、观测个案),以提高统计推理的信度,我们从差异性来考虑,如果增加的个案而未增加差异性时,增加样本单位并不能对提高统计推断的信度作出任何贡献。不难理解,如果实验条件绝对一致,观测个案完全同一,那么一千次实验,一千个个案并不比一项实验,一个个案更有力量。

增加差异性就增加样本与总体的同构性,当观测的样本单位的差异性已经再现了总体各单位的差异性,就是完全同构也即达到了“完全相似”,那么就做到了对实验的完全控制,或对总体的完全观测,这时,再增加样本单位数也就没有什么意义了。可见,增加样本容量的真实直接目的,是为了扩大差异性,接近样本与总体的同构,在“全正相似”,或已经实现了“全负相似”的情形下,再扩大样本容量都是无意义的。通常所说的增加样本单位以提高可信度是基于以下背景:(1)总体全正相似,即对所有总体单位来说有性质 A,那么它也具有性质 B,这种情形在实际中是不需进行统计研究的。因为,没有差异性(无差异的总体在实际中几乎是不存在的)就无须统计。(2)全负相似是一种理想境界,在实际中几乎是达不到的。因为,我们不可能准确地知道总体各单位在所有方面的区别方式,因而也就不可能穷举所有不同组成部分的样本单位,或穷举所有实验条件。(3)一般说来,不会有两个总体单位是完全相同的,因此,增加新的样本单位就总有可能引进新的“负相似”。从而加强统计推理的可信度。不断增加样本单位数,不断接近“全负相似”,即同构。

增加样本容量还有“净化”相似,从而提高统计推理可信度的作用。这里的“净化”相似是指使样本单位之间的相似唯一化。例如,在  $n$  个样本单位中具有 A 与 B 的正相似,即具有性

质  $A$ , 则就有性质  $B$ 。同时, 还存在另外一个性质  $C$ 。也属于该样本的正相似。统计研究的目的是在于建立相似性  $A$  与  $B$  之间的本质联系, 得出只要有性质  $A$  就有性质  $B$  的结论。但在现有样本中还存在另外的相似性  $C$ , 那么不禁要问,  $C$  与  $B$  之间是否存在一定的内在联系呢? 甚至  $A$  与  $B$  关系是否是由于  $C$  与  $B$  的关系引进的呢? 这时, 统计研究者的第一个想法就是设法破坏  $C$  与  $B$  的关系, 若增加一个样本单位, 该单位具有性质  $A$  与  $B$ , 但不具有性质  $C$ , 这样在  $n+1$  个单位的样本中,  $A$  继续保持着正相似,  $B$  也继续保持着正相似, 而  $c$  由在  $n$  个单位的样本中的正相似转变为在  $n+1$  个单位的样本中的负相似, 从而破坏了  $C$  与  $B$  的联系, 使得样本中留下的只是  $A$  的正相似和  $B$  的正相似, 我们说相似性得到净化, 排除了  $C$  可能与  $B$  有存在联系的疑虑。

显然, 增加样本容量的直接目的, 仍是增加差异性, 使之接近全负相似。

最后, 当我们没有足够的理论或事实依据证明在所观测的个案中确实存在有  $A$  则有  $B$  的关系时, 比如说在  $n$  个个案中有部分有  $A$ , 则不具有  $B$  的个案, 这时, 一方面需要进一步掌握有关科学理论知识, 另一方面就只能靠增加观测个案, 来调整我们关于已知性质相互关系的认识。

人类认识就是寻找客体的真值, 统计方法提供了一种无限逼近真值的工具, 只要真值确实存在, 人们就可根据不断修正的方法无限逼近。这样, 世界是可知的这一认识观在统计认识过程中再一次得到证明。

## 二、参数估计

由样本信息估计总体参数这是归纳, 但是怎样根据样本求总体参数, 参数函数如何构造, 又大大不同于“归纳概括”。

科学哲学中有如下一致见解: 并不存在解释特定事实的唯一确定的、具有绝对真理性的科学理论, 建立别的理论作重新解

释总是可能的,问题是在不同的理论之间进行评价,比较它们的优劣。在统计认识中,不能认为待估参数  $Q$  有一个绝对的、先验的值,而其它的值都只是对它的近似。应该看作所有的估计值都是  $Q$  的一个确定的值,而又都不是唯一有资格的值,不同的估计值客观有优劣之分,这取决于估计方法的构成和样本的质量。因而关于统计方法的比较和评价则显得格外重要。对估计方法进行评价、比较的一系列判据,如无偏性、有效性、一致性等,再一次说明就统计估计来说,由个别观测值到表征总体的参数值,是一个归纳过程,都不同于通常理解的种种归纳,它需要执行一套复杂的专门技术程序;另一方面对这套程序进行评价,或对可能出现的不同程序进行比较,就象科学哲学中理论的评价和比较一样,是有章可循的。

### 三、假设检验

统计假设是以已知的原始信息和科学原理为依据而对未知的参数所作出的假说性说明。它是一种探索性的认识形式,是试图揭示事物内部机理和深层本质的认识活动。从认识论上说,事物的本质“隐藏”在现象背后,无法直接观测,只有通过假定性说明步步逼近。这就决定了统计假设的核心内容是关于客体本质的假定。而这种假定是基于一定事实基础和关于客体的有关理论提出的。这里的事实基础可以是现实的历史的事实,也可以是经验,甚至可以是理性猜测或相似系统的相关特征值。统计假设应该能推出预言或预见,以便进一步验证。因此,完整形态的统计假设应包括如下结构性要素:事实基础;本质性假定;预言预见;有关理论(包括推理规则)。

在这些结构性要素中,事实基础是提出假设的基本依据,仅在出现确凿的而又被已有理论满意解释的事实时,提出假设才有必要,假设在验证前的科学性应由事实基础保证;本质性假定是假设的核心,假设的认识价值集中表在这里。假设的验证也主要针对这部分进行。预言预见是由本质性假定演绎出的未

知事实,它提供对假设进行验证后评价的可能。背景理论是假设中暗含的内容,包括被假设借用或修正的各种理论知识。

统计假设的上述构成决定了它具有两方面的特征:一是它具有一定的科学依据,二是它具有一定的假定性、猜测性。任何统计假设都是在一定的原始信息和科学理论的基础上提出来的,它能够解释与它有关的现象,并与它引为根据的已有理论相一致。不同于无根据的猜测和其它神话、幻想。但是,统计假设的核心内容毕竟是推测出来的,尚未得到验证,因而它又不同于成熟的科学理论,统计假设正是科学性和假定性的对立统一,它向科学理论的发展过程就是假定性成分得到实践确证,从而与科学性统一为客观真理的过程。

因此,统计主体必须对其所作的假设进行检验。即按照统计规则检验统计假设是否符合统计事实。在统计假设中, $H_0$ 总是一个关于“无效应”(no effect)的陈述, $H_1$ 是对 $H_0$ 的否定。统计假设检验并不保证我们所接受的假设绝对可靠,只告诉我们,在已有的知识背景和证据上怎样做才是合理的。检验遵守陪审制度下的被告原则,即 $H_0$ 相当于法庭被告,除非现有的事实证明它有罪,它便无罪,除非处理结果强迫我们接受 $H_1$ ,否则便接受 $H_0$ 。统计认为,在已有的知识背景上,小概率事件可以认为实际上不会发生,如果居然发生了,那就表明另有未知的因素在起作用,因而需要对背景知识作出增补和调整。可见,统计认识至少起了思维定向和提出问题的作用。

总之,统计假设告诉我们在已知的知识背景下,怎样做才是合理的;统计处理方法并不遵从任何人为的禁令;统计观点决不排斥和取代任何别的可能的观点;统计结论并不告诉我们假设的源缘、内容、机制等。

#### 四、统计比较

比较和分类是根据事物之间的异同进行分门别类的研究的一种方法。比较可以确定对象之间的异同点。它首先是把客体

的各方面的信息予以分解(析),然后在有选择(分析)的组合(综合)的过程中判明某些信息的异同。可见,比较是从分析开始,以综合终结的。分类可以是比较过程的后续。它在比较的基础上进行求同或求异的分类。在求同中,它是某种信息的综合;在求异中,它是某些信息的分解。比较和分类其实是在分解和组合的过程中对认识对象之间诸方面的信息进行别异、求同的一个具体方法,在分解和别异时,它是分析;在组合和求同时它是综合。

### 五、模型外推

模型外推是统计模型方法的一部分。统计模型研究有三大主题:一是构造符合相似条件的模型,即再现原型;二是对模型的实验研究,在实验和测量条件下观测它的行为,用相应的方式加以记录;三是把对模型进行实验研究所获得的结果中得到的信息外推(转移)到实在的被研究的客体上去。前两个主题属于统计描述,而第三个则属于统计推断即模型外推所研究的内容。模型外推法的基础是模型和被模拟客体之间的关系。前已述及,统计模型是这样一种实验研究方法,它以模型代替所研究的客体为基础,而模型本身与客体(原型)处在客观相似关系之中,这种相似作为几何的、物理的、数学的相似关系而能够被精确化、形式化。外推的合理性由相似理论来决定。相似理论的第一条定理,表明只能转移到用同样方程式描述的现象上去。第二条定理表达出对材料的一定加工方式的条件。它要求确定的不是个别量之间的依赖性,而是作为相似标准的综合体之间的依赖性。第三条定理给出现象之间相似的主要条件。

统计模型外推的数学方程式,描述的是具有不同的物理本质,但具有类似结构(数学模型、物理类似物)的系统的相同行为,以及执行着相似的功能且以自身成分的物理本质及其结构为特征系统的行为。

模型外推法按其性质和形式来说,它接近演绎推论,因为信

息的转移是按建立在使用相似标准的基础上的规则而实现的。同时,它属于认识的经验层次,因为构造模型的步骤、对模型的实验,对它的行为的观测和测量都是实验的、感性的、对象——工具的活动。因此,对统计模型的推断结论在信念上和应用上都不应过分,只是一种探索,是否是真理、科学,还需要检验、证明。

对所得到的材料以标准方程的形式进行加工和按照相似理论的规则把信息转移到客体上是一个演绎的过程。因此,模型外推法本身是演绎推论的特殊形式,而模拟法作为整体来说是归纳和演绎的统一,是科学认识的经验方法和理论方法的结合。

### 3.4 统计信息利用方法

#### 3.4.1 统计解释

##### 一、统计解释

统计解释就是统计认识主体根据统计认识结果对客体作出合乎规律性的阐明。即把客体涵盖在某种规律起作用的范围內,从一般性的规律中说明个别现象为什么会发生以及可能发生的状态及其概率。解释客体意味着揭示它的本质,因此,统计解释在于揭示客体客观存在的“很重要的”联系和关系。要回答客体是由那些主要成分构成的,各要素是怎样相互联系起来的,要素对客体的影响方向和程度,要素之间的相互关系、客体的变化规律等等。统计解释可以分为:

1. 状态解释:揭示客体在不同历史条件下的形态及其发展进程,包括所达到的总规模、总水平、一般水平和差异程度,以及分布规律等。

2. 结构解释:在于揭示系统的组成元素、各元素之间的联系、各元素的重要性(如构成比例、对客体的影响程度等),以及

用结构决定系统的某种属性、行为或结果。

3. 因果解释:探索影响客体变化的因素以及各影响因素的作用方向和作用程度

4. 功能解释:阐明客体系统对其环境之间的相互关系和相互作用。

5. 规律解释:这里的规律是指统计规律,一方面给出客体可能存在的规律,另一方面又给出这种规律的概率。这里对概率的不确定性应该有一个正确的认识。从本质上来说,不是我们所获得的规律具有不确定性,即知识的不确定性,而是规律本身是不确定的,其不确定性属于客体本身,是客体的非单值决定论的行为客观决定的。当然,获得的知识也存在不确定性,毕竟是一次观测的结果,是现有信息的产品,也就是说,统计认识结果仅在进行统计认识研究所依据的信息范围内是正确的。超出这个范围以外就不能保证其正确性了。需要不断获取信息逐步修正和完善。

6. 趋势解释:描述客体的历史变化轨迹、发展变化样式、速度、以及未来走势等。

统计解释中,如果被说明的现象不是已观测的现象而是待认识的现象,那么这种解释就属于预言或预见。因此,预言实际上是对有待认识的现象的解释。

作出解释,意味着需解决四类问题:(1)进行解释所依据的统计规律是什么;(2)规律起作用的条件及范围是什么;(3)被解释的现象是什么;(4)如何建立上述因素之间的推理链条。这几类问题均需在统计认识过程中解决。因而,统计认识就表现为不同的类型:(1)规律发现(概念、定律、理论、相关关系、统计规律、质量互变界限(度)、发展趋势发现);(2)条件和现象发现(状态、结构、因果关系发现);(3)推演形式发现(方法、程序、操作规则发现)。

解释的逻辑结构是指两种判断之间的关系。一种判断中表

达出来的是理论的解释性论断,另一种判断中描述的是被解释的事实。前者常称之为辩解,后者则为注解。辩解与注解之间的逻辑关系的特点,决定了解释的逻辑结构可能是演绎还是归纳。当辩解中的解释性前提是真理性论断时,为理论解释;当辩解中的解释性前提是假说时,为假设性解释。前者得出的结论是逻辑必然,只要逻辑结论符合对所要解释的事实的描述,就是真实的和可靠的解释。而后者得到的逻辑必然的推论是可能的,是概率性解释。在概率性解释中,对个别事实解释的概率性与“统计规律在每一个别情况下无法实现”这一规律联系着。因为统计规律是大数规律,它仅在大量观测或多次试验中方能出现。因此,在统计规律基础上所作的关于个别事实的结论,只能解释这一事实的可能性,而不是它的必然性。

在统计解释中,最主要、也是最为重要的解释是统计模型解释。统计模型是对客体作出科学解释的依据,它是对客体本质和运行规律的系统化反映,具体地刻画了客体的本质特征及客观存在的关系。

对模型解释的批判:模型解释实质是把从一类现象中获得的理论解释转移到新的现象领域。这种转移的依据是两个领域间的相似或类比关系。但是,无论模型如何准确,类比关系的存在一方面推进发现,帮助想象到解释性假说,因而有启发作用。而另一方面,也能成为错误的根源。因为任何类比都是由三类关系组成的:被比较的现象和它们的模型的相同特征之间的关系(“肯定的类比”);它们的不同特征之间的关系(“否定的类比”);不清楚它们之间的关系是相似还是不相似的那些关系。判明这些关系本身就是将假说性解释转化为真实的理论性解释的发现。因此,在还没有作出发现和确立理论以前,借助于模型将对一个领域的现象的解释转移到另一个领域的现象上是不完全的,仅是局部的,只是在被比较领域的规律性相似的范围之内才是可行的。在统计模型分析中,选择模型时,判明所研究的客体



在已有理论上和所采集的数据是否具有模型的基本特征,是否满足该模型的基本假定,实质就是在判明相似性。换句话说,统计模型就是建立在这个相似关系的基础之上的。因此,这是非常重要的一个环节。由此看来,统计模型解释从方法理论上来看,只能被看作通向可靠的、真实的和在理论上更加适合的解释发展的路标和阶段。

借助于模型的解释会把模型的某些对于所研究客体来说是不存在的特征一同转移到所研究客体之上的危险,因而对在模型中所体现的理论的运用有可能走到无成效的、错误的方向上去。因而在应用模型解释客体时,必须与客体的理论紧密联系,将定量分析与定性分析有机地结合起来运用。

因此,模型解释就其本身的逻辑特点而言,属于非演绎判断,同时就其认识论本质而言,它既不同于理论,也不同于假说。

## 二、统计解释的验证

在统计认识中,解释是一个基本环节。然而解释有真假之别。正确的解释需由真实的规律、准确的条件和结果以及合适的推理形式构成。只要上述成分中有虚假因素,解释的真理性就难以保证,因此,必须对解释中所作出的统计认识进行验证,才能排除各种假解释,这就使验证成为统计认识中另一个重要环节,各种类型的发现均需通过验证以判定其真伪。

验证,是指对各类统计认识进行真理性评价的过程及其结果。按照实际用以验证的评价尺度,可将验证分为三类:一是逻辑判定,其判据为客体内在的一切逻辑关系;二是历史验证,即统计认识结果是否与客体的历史事实同一;三是实践检验,就是将统计推断结果与实际结果相比较。各种验证相互补充,可以结合运用。显然,验证与解释是统计认识中两类不同性质的问题。解释是求解过程,而验证则对求解结果进行真假评价;验证是解释的延伸,解释中就应当包含着验证的可能性。

### 3.4.2 统计预测

统计预测是统计主体根据客体发展的一般规律,对客体未来的发展态势以及对客体系统环境等影响进行预言和推测。统计预测的科学性在于这是对客体发展的具体条件表达的结果,是客体内在存在的数量变化的表述,是客体本质的数量逻辑推断,即有理论依据、符合客体本质、遵循逻辑规则、精确计算的结果。

统计预测是统计认识的创造性能动作用的最突出的表现之一。显示出统计认识能够明显地、合逻辑地超越一般认识的经验层次。

在统计预测中如何根据现实去认识未来是非常重要的。根据当代预测的实践,人们总结了一些经验,并表述为若干中介性的基本原理:

1. 惯性原理(统计动态预测的理论依据)。惯性是客体依时间发展的特性。从时间角度来看,现在是过去的延续,未来是现在的延续。如果一个客体现在的行为是过去行为的延续,则称该客体具有惯性。对于一个客体来说,如果客体机理没有发生明显的变化且也没有受到环境的大的干扰,那么,客体将保持自身原有的运动状态;当客体机理发生显著变化或受到外来强大干扰的影响时,客体按一定的记忆方式,保持着对该干预或扰动的记忆。也就是说,干预对客体的影响力与进入的时间点无关,仅与时间间隔有关。任何客体不存在有无惯性之说,只有惯性大小之别。惯性越大,过去和现在对未来的影响就越大。所以,统计主体可以对客体的惯性轨道进行精确的测算,并根据惯性原理进行统计预测。即当客体系统处于相对稳定状态,尚未产生突变的连续过程的时候,只要人们了解了它的过去、现在的形态和规律,也就近似地认识了它的未来最可能的形态和规律。

2. 相关原理(统计回归预测的理论依据)。事物的发展不是

孤立的,而是与其他事物处于相互联系、相互制约之中。当某个系统中的一个事件发生变化,其他事件也会受到或大或小的影响,这就是相关性。相关性可表现为多种形式,其中最常见的是因果相关。人们只要知道了过去和现在的原因,就可以推测出将要产生的未来结果。利用相关原理进行预测,首先找到与被预测事物具有较强的相关性且其发展变化趋势又已确定的另一个或另几个事物;然后还要找出这些事物与被预测事物的相关程度。关键在于相关性的<sub>大小</sub>,若相关性较小,就不宜按相关原理进行预测。

3. 类推原理(由类似条件推类似结果,如领先指标预测法等)。从历史上出现的事例中,找出它们产生的共同条件,如果在未来的发展中,还会再具备这些条件,那么历史情况可能重演。类推的前提是寻找相关类似条件。如果被认识客体与某一现象的发展变化在时序上有差别,而在发展的表现形式上有相似之处,那么就有可能把先发展的事物之发展过程类推到后发展的事物上去,从而对后发展事物的前景进行预测,这就是类推原理。运用类推原理时,一定要在关注相类似条件的同时又注意条件变化的差异性,以提高预测的准确性。先行指标预测就是一种类推。

根据不同的标准,可以把统计预测方法分为不同的类型,一般有四种基本类型:

一是直觉型预测方法。即主要依靠有关专家的经验、知识及其直觉的分析和判断能力进行预测。

二是模型预测方法。即主要根据预测对象的时间与空间、定性与定量、过去与现在的资料、数据或状态,建立相应的统计模型去逻辑地推测其未来发展的状态和趋势。

三是类比性预测方法。如果在两个系统之间具有相同或相似的特征,已知其中一个客体系统的发展变化过程,根据类推原理,就可以类比预测另一个客体系统的发展趋势。

四是演绎性预测方法。根据有关预测对象的历史和现状的资料和数据,选取一个恰当的数学模型,运用数学方法求解所选预测模型的待定系数,从而得到一条表示预测对象发展趋势的曲线,据此,进行外推就可以得到预测对象未来发展的技术特征。

统计推断未知和预测未来只能是确定客体未来发展的基本方向、总的趋势,而不可能对其变化的具体细节作出推测。因为客体的变化和发展,不是在理论空间运动,而是存在于现实世界。因此,其变化和发展是一种极为复杂的矛盾运动,会受到各种各样因素的影响,且影响方式和作用程度也不完全确定,况且在未知的时空域还会出现新的影响因素,生产新的关系。所以统计推断和预测是在一定概率条件下进行的,是一种整体思维、大众思维、或然思维。换句话说,它是立足于现实条件基础上,提出未来各种可能性(未知状态),形成一个可能性空间和或然性空间,统计认识主体根据自己的认知结构、制约条件、价值取向,在可能状态空间中进行择优。可见,现实条件和制约条件及认识手段本身决定了统计推断和预测的客观性,但是认知结构,手段选择及价值取向又体现了统计推断和预测的主体能动性,是二者的统一。

预测与解释既有联系又有区别。联系在于都是依据同样的道理(即同一些规律性和本质性联系),并利用同样的逻辑“机制”从原始的普遍理论前提下演绎出结果。即:(1)依据的理论相同,(2)运用的逻辑机制一样。区别在于:(1)解释是对从前已知的事实的解释过程中寻找理论,由这个理论逻辑上推导出在内容上同对这些事实的数量描述相适合的判断;预测则是从该理论中逻辑地推导出描述还不知道的事实(属性、现象)的结论,也就是在未能和应当产生的那些事件,和已经存在、尚未认识、应当在将来被发现的那些事实。(2)解释提供了认识过去和现在,揭示已知事实的本质和从理论上领悟它们的可能性,预测

使统计认识主体可能认识事件的进程和趋势、前景。

### 3.4.3 统计决策咨询

统计决策认识就是根据统计推断认识所揭示的客体的结构和运动规律,结合宿体的目标和各影响因素的状态及概率,拟定出若干可行的行动方案,并按照一定的准则从中选择最优方案,以指导实践,有效地达到预期的目标。统计决策认识过程是价值理性认识的过程,它经历了确定目标,拟制方案,评价选择方案等程序,是思维从确定目标到做出决定的一个动态过程,是一种创造性思维。

鉴于刘李胜博士在其所著的《决策认识论导论》一书中,对决策咨询方法问题进行了较为系统、深刻的哲学研究,也考虑到本文的篇幅所限,故在此暂不作进一步研究。

# 4 统计思维

## 4.1 统计思维的特点

### 4.1.1 统计思维一般

统计思维就是统计认识主体在其认识过程中,运用统计方法、统计概念、统计知识思考客体的一种较为稳定的致思取向和思维特征。它以运用统计方法作为核心要素和最重要的表现。统计思维不仅贯穿于统计认识主体处理统计认识问题的思维活动中,而且也体现在其他科学工作者和管理工作者处理认识问题的思维过程中。统计思维在其实际的运作过程中,常常经过以下步骤:首先根据宿体的认识目的,在统计理论和有关客体的实质性科学理论的指导下,对所认识的客体进行深入分析,找出需要且可以测量的本质特征;然后,确定各个特征的内涵、外延、量的规定性,使其形成为一个统计指标即转化为一个统计变量,进而根据各个统

计指标之间的内在关系建构起统计认识指标体系,从而把所要研究的问题简化为一个统计问题;最后,按照统计规则,运用统计方法进行信息采集和加工,实现对客体特征和规律的统计把握。

统计思维的特征是:不论任何客体,不论任何目的,都首先将所要研究的问题抽象为一个统计问题,将所研究的客体系统转换为一个信息系统;然后,按照统计认识程序,运用统计方法进行统计信息加工,得出统计认识结果。统计思维是人类认识世界的重要思维方式,其优点是客观、全面、严密、精确、系统、深刻,易于主体在定性与定量的辨证统一中准确地把握客体。

统计思维的基本规律是“认识同一律”即统计认识结果必须与客体的本质特征相一致。所谓客体的本质特征,从哲学的角度看,就是客体的质的规定性、量的规定性和质与量的统一。从思维学的观点来看,统计思维同样也是对客体深远区层实现穿透性反映的物质运动。对于反映主体来说,客体分为表近区层对象和深远区层对象;可直接感知的是表层对象,可直接接触的是近区对象;不可直接感知的是深层对象,不可直接接触的是远区对象(又有空间上的远区和时间上的远区之分)。对于表近区层对象的反映是直接摄取方式,对于深远区层的反映则是穿透复制方式。穿透就是透过表层捕捉到深层的信息,通过近区转录到远区的信息;复制则是把深远区层的信息还原为对象的实在状态或真实映象。主体对客体深远区层认识的深度和精度标志着主体对客体的把握程度和可能的控制和利用程度。因此,客观、精确地把握客体深远区层的特征是统计主体和客体永远追求的目标。但是,对客体深远区层的认识,经常出现认识结果与实际状态不同一的现象。这种不同一往往导致认识无法继续深入进行下去,甚至出现误导。所以,在统计思维中,特别要求统计认识结果必须与客体深远区层的实际状态相一致。

统计主体对客体深远区层反映结果与实在状态经常、大量

地不同一,而统计思维的客观规律又要求必须同一,这就是统计思维的基本矛盾。

统计思维的工具就是统计逻辑。即统计主体进行统计认识的思维形式、方法及其规律(具体表现为各种统计逻辑规则)的学说。作为统计思维工具的统计逻辑是统计认识的总结,是人类所得到的对世界的统计认识的历史的总计、总和、结论。同时,统计思维作为探求新知识的方法,不断地重新回到具体的统计认识过程中被精炼、概括、上升为对客观世界进行再认识的方法和规则。统计逻辑是一种应用逻辑,是具有从量的方面经过信息测量和加工,观念地把握和认识客体的总体特征和规律性的方法意义的普遍知识。

概括地说,统计思维就是统计主体根据理论、经验及观测数据形成对客体特性的统计认识过程。因此,统计思维是人类社会一定阶段的产物,人类历史上每个时代有什么样的思维方式,就相应地有什么样的统计认识模式。现代统计思维方式具有社会化大生产的整体性、开放性、有序性和动态性等特点。要求人们以辩证唯物主义的系统观为指导,以信息处理技术为依托,按照统计的逻辑推理,结合系统的内部结构和外部环境,开阔思路、精确分析、勇于探索,使描述方法和推理方法建立在更加科学和更加严密的基础上。

#### 4.1.2 经验思维与理论思维的统一

统计认识要搜集和处理浩繁的统计数据,以数据(信息)为材料进行认识活动,把感性认识上升到理性认识的高度。统计认识过程不仅是通常所说的实证性研究活动,同时也是探索性研究活动。它自始至终都是理性认识和感性材料的相互结合和相互渗透。因此,统计认识作为一种思维过程是以理论思维为主导,具有两种思维的成份,各种思维相互交叉,相互补充,使统计认识更系统、更具体和更深刻。



按照统计认识属于实证性研究来说,它具有经验思维的特点。经验思维就是运用实践经验、感性认识和感性材料进行的思维活动。它的功能主要是认识具体事物的外部状况、表面联系和现象,通过经验思维能够对丰富的大量材料初步加工,把握事物多种多样的具体状态,并且能够在一定程度上把握事物的内部联系和规律。统计描述就是一种比较典型的经验思维。它依据的是客体的个体的实际状况或者是客体过去的、现在的状态,是事实的归纳、概括、整理。例如。我们通过对整体的各个(或足够多的)个体的特征进行观测,得出关于总体的总量,构成等特征;通过对客体过去的轨迹及现在的状态,标绘出其一般发展变动趋势。这些认识均属于表象、经验。客体系统客观上是否具有这种机理,有待进一步进行理性分析研究。

从统计推断来看,它在统计描述提供的经验材料的基础上,运用一定的理论、概念、依据严密的逻辑规则和推理过程进行假设检验、数理推断、悖论分析,对描述信息、经验认识进行理论思考,使经验认识升华,这又具有理性思维的特点。它抽象掉具体个体数量上的差异,得出有关对象的共同本质特征的认识;抽象掉所依据的经验材料的特殊,得出有关“类”的一般的认识。例如,在研究父亲与儿子的身高时,所运用的观测数据是一定容量的样本,而得出的结论却是对人类一般的认识;同样关于某一地区男性人口的寿命,也不是具体某个男性的寿命,而是该地区男性的一般寿命。

我们说统计信息来自于具体而非具体,来自于实践而非实践,来自于特殊而非特殊。这就是说,从物理(physical)的(实体的)角度来看,“白马非马”,但是,统计是对信息的考察和研究,恰恰就是由“白马”到“马”。

#### 4.1.3 直觉思维与抽象思维的统一

按照统计认识要运用材料来看,统计思维又具有直觉思维

的特点。比较一般地认为直觉思维依据一定的数据和事实,使人们得到启发,运用已有的经验知识,对客观事物的本质及其规律性作出迅速的识别和直接的理解,并对对象的总体状况作出判断。统计认识在取得统计数据之后,首先就是根据数据的特点,运用一定的数据整理手段(如分组、直方图、茎叶图、频率图等)和统计研究人员积累的统计认识经验,充分发挥主体的能动性,获取初步认识。在此基础上再对统计数据的背景资料进行分析研究,必要时还要进行典型剖析或抽样验证。这一阶段的统计思维具有明显的直觉思维的特性,属于直觉思维的范畴。虽然统计直觉思维的认识结果缺乏充分的理论依据,得到的认识可能肤浅、粗枝大叶,甚至有时是假象(如:数据本身具有某种特征或是受随机影响,代表性不强等非所研究系统所具有该种特性),但是,统计直觉思维不仅是抽象思维的基础,而且直觉思维是最能充分发挥统计认识主体的主观能动性,是最活跃、最遵循客观事实、最富有创新性、最能获得独到见解的思维过程。到了理性思维阶段,统计研究者必须遵循严密的逻辑规则和一定的推理过程。事实上,当统计认识的数据、推理方法和计算手段确定之后,不论统计研究者是谁,是否具有丰富的经验,都会得出同样的结果。在某种意义上说,这一阶段被程式化了。

由于统计直觉思维缺乏一定的理论依据,因而其认识结果的可靠性和可信性不强,同时,由于直觉思维是就事论“事”而不是就事论“是”。这里的“事”指事实,而“是”指规律。“事”是个别、具体、特殊,而“是”是共性、抽象、一般。因而统计直接思维不能得到“类”的一般性的认识。

在一定理论指导下进行的数理研究,又具有抽象思维的特点。属于抽象思维的范畴,它舍弃具体向客体的规客规律性逼近。

#### 4.1.4 形象思维与逻辑思维的统一

从统计认识的表现手段来看,统计思维具有形象思维的特

点。形象思维是通过感性形象,如形状、色调、图案等,揭示、理解、认识客体的状况。统计认识常常利用数据图、分布图、趋势图、结构图、相关图、象形图等手段,通过图形的形状如条形、柱形、饼形、球形、象形,颜色等来揭示和描述认识客体的状态、构成、发展变化情况属属性。使客体的状态、结构、变化及与其它现象的相互关系活灵活现,既直观、易于理解,又有助于加深印象并激发认识主体及宿体的灵感,客观、准确、全面地把握客体。但是,统计认识决不以形象思维为终结。恰恰相反,形象思维仅仅是统计认识的开始,旨在给统计认识主体以启发,提示客体可能具有的属性和存在的规律以及进一步认识研究的方向。在此基础上,按照一定的逻辑规则,经过统计推理,得出逻辑结论,即进行统计逻辑思维。统计的逻辑思维也是将概念、判断和推理三种基本形式有机地结合在一起。统计认识一般可概括为从现有的个别的概念(信息、指标)出发,经过统计判断,实现由一种信息向另一种信息的转化。一般认为统计是一种推理,这是不全面的。每一次统计认识活动,都要经过由概念、判断和推理的相互转化。判断确定各个个体是否具有某种特性,例如在了解某一种药物的疗效时,首先要根据这种药物的主治功能、用法用量、服用对象以及用后的主要疗效等概念,对被观察者上述各个方面的、个别的、具体的特征进行判断,然后在此基础上经过统计加工、综合、“浓缩”、推断,得出反映普遍的、一般的、本质的特性——统计指标来。缺少任何一个因素,中断任何一个环节都不能完成统计认识。

#### 4.1.5 归纳与演绎的统一

统计认识是通过个别研究认识一般的,所以统计思维必然是一种归纳(即必须通过归纳才能实现)。正如亚里士多德所说的“要认识一般,不通过归纳是不可能的”(亚里士多德:《后分析篇》,见《工具论》,广东人民出版社,1984年,第191页)亚里士多

德认为:没有归纳,就不可能从一般中“取得的”知识;没有论证,就不能得到具有必然性知识的结论,统计不仅要根据所构建的原始信息通过统计推理获得一般的“知识”,而且还必须进行假设检验、机理检验等,对所获得的知识进行论证。所以说,统计思维是归纳与演绎的统一。归纳方法论强调了方法和外来信息的重要性,而演绎方法论则强调了问题和先存知识的重要性。实际上,二者是一个有机的整体,需要相互补充和协调才能真正解决问题。统计学是认识方法论,既重视未来信息和方法,又重视问题和先存知识,认为统计思维只是归纳的观点是完全错误的。如回归分析既是归纳,又是演绎。比如说,在成本对产量的回归中,一方面从条件相同的各个不同企业的产量与成本的对应关系出发,通过统计归纳,抽象掉个别差异,得到该类企业成本与产量的一般关系式: $Y = A_1 + B_1 X$ 。另一方面,在进行归纳时,就是从先前已经得到的规律  $Y = A + BX$  着手的,是把一般的  $A$ 、 $B$  演绎为个别的  $A_1$ 、 $B_1$ 。所以说,统计思维将归纳和演绎高度而有效地结合运用,收到了很好的认识效果。

#### 4.1.6 探索与实证的统一

从逻辑的角度考察,统计学旨在建立和完善能获得认识对象的有关知识的科学理论的思维工具。思维无非就是按照对象本身的形式和尺度,根据客观上正确反映该对象的映象自觉作用于任何对象这种能力的发展。这表现为具有探索、获得认识的功能,同时也具有实证的功能。统计认识把发现规律、证明规律有效地结合起来,并强调探索和实证都必须从事实出发,按客体所固有的规律进行。获得新认识同证明认识有着不可分割的联系,并且证明认识是获得认识过程的从属环节。为了证明任何一个认识结论的真理性,就必须揭示我们的思维通向这种真理性的途径,就必须分析数据资料,分析加工数据资料的规律和方法,分析理论的构成方法。发现与证明二者是相互联系、相辅

相成的一个整体,不是先发现,尔后才证明的割裂状态。发现认识过程本身就含有对认识的证明;反之,对认识的证明同时也就是对认识的深化、精化和具体化。

认为统计实验只是证明认识结论的真理性的工具或仅仅是发现新规律、构造新假设的手段,这是不正确的。统计实验和其它一切科学实验一样,是获得新的认识结论和证明或反驳某个现有认识结论的统一。当我们提出某种新的认识结论时,我们同时也就反驳某种陈旧的认识结论和证明某种新的认识结论。除了判断统计认识或其它理论、假设的真伪性外,证明过程没有任何别的目的;反之,获得统计认识的过程也把证明作为一个环节包含在自身之中。

在这种意义上说,统计学和哲学一样既是研究的方法,即探索、获得新认识、新规律性的方法,同时又是证明现有认识和规律的方法。探索规律的过程也把实证分析作为一个环节包含在探索研究之中,实证分析的过程同时也是肯定新认识、否定伪假设的过程。

#### 4.1.7 实践与理论的统一

不仅作为获得认识的统计推断要从对象的“自己运动”,从对象的一切联系和“中介”出发研究对象,而且作为证明认识的统计检验也要如此。实践在获得认识和证明中具有特殊的意义。统计观测,统计实验,统计抽样都是对实践过程的研究,其结果是实践活动的客观、真实记录。培根曾在《新工具》中把实践归结为观察和实验。离开实践根本不可能解决任何一个认识结论的真假,理论和实践的统一是马克思主义哲学最主要的方法论原理,是研究对象和判明已得知识真理性的一根红线。我们知道,如果一种理论,一个命题,一个假说是通过逻辑途径从另一些真实性早已判明了的原理中推出来的,那么事实上它就被认为已经得到了证明。然而,如果不超出思维范围进入实践

活动领域,那就既不能解决在证明中作为推断依据(论据)的任何一个科学原理的真实性问题和数据信息的客观性、同构性问题,也不能解决推断方法的相合性即统计方法选择的正确性问题。我们的推断究竟是客观的即统计信息同被认识对象本身同构完全蕴含着对象本身的结构、属性和规律呢,还是统计推断已经脱离了被认识对象所固有的属性和规律性而在数理逻辑领域中活动呢?统计推断所运用的统计方法是相合的即统计方法的基本假设(应用条件,范围等)与统计信息集和对象本身属性相符合、相一致呢,还是统计方法早已违背了基本假设而在数学领域中运算呢?如果忽视实践在证明中的作用,那就无法回答这些问题。

#### 4.1.8 历史的与逻辑的统一

动态分析方法是典型的历史的和逻辑的方法的有机化合。历史的方法是按照事物产生和发展的实际进程,对依次出现的事实进行系统的分析,发现历史发展过程的基本线索,揭示事物的本质和发展规律。逻辑的方法就是从典型的发展阶段(或事实)入手进行理论研究,以概念、范畴的理论体系,揭示事物的本质和规律。其显著特点是抓住客观事物的本质联系和必然趋势,而不完全局限于客观事物的历史本身。撇开了哪些偶然的、支流的、和暂时的因素,集中地反映客观事物历史发展的规律性。

编制时间数列、利用增长速度、发展速度、发展水平、增长量等方法,是一种简单、一般化的历史分析方法,可以说是“就事论事”,往往会受到许多偶然的、暂时的因素的影响。苛刻一点讲,这些方法只是对认识客体历史状态(行为)的记录和描述,没有探究其本质性的特征,因而达不到对规律性的认识和把握。因此,与其说它是一种分析方法,倒不如说是描述方法。时间数列的传统分解方法、趋势分析方法、随机时序模型方法等,则是“就

事论是”，它抽象掉了偶然因素的影响，撇开了支流和次要因素的影响，集中揭示客体的本质特征、基本趋势、一般规律。因此，勿宁说它是一种动态分析或历史分析，倒不如说是历史的和逻辑的有机化合。

## 4.2 统计推理范式

### 4.2.1 基本范式

统计思维在推理过程中，就其探索性来看，往往是一种归纳推理。但是，许多推理过程，包括进行归纳推理的依据以及对统计归纳推理的结论的检验等都表现为由一般性前提推出个别性结论。也就是说符合演绎逻辑。不过不是纯粹意义上的演绎推理，它是根据一般性的前提判断依统计规则作出个别性的结论判断。显然，它与一般意义上的演绎推理的区别就表现在“依统计规则”这一点上。这里的“依统计规则”表明了两个方面的含义：一是前提与结论的关系不是必然蕴涵，而是统计蕴涵；二是由前提推出的结论不是必然结论，而是统计结论。所以我们称之为统计三段论。此外，需要特别强调的是这里的“三段论”也不是仅指演绎而言，而是强调统计推断过程中由前提到结论的逻辑，并不限定前提是一般性的，还是个别性的。也就是说，统计三段论也包含统计归纳推理的形式在内。统计三段论在统计推理的不同阶段具有不同的形式，在这里我们不进行全面研究和系统概括，仅讨论几个基本范式。

范式一：

根据统计法则得到的信息在总体上与客体同构

原始信息采集符合统计法则

原始信息在总体上与客体同构

原始信息与客体同构这是统计认识的基本假定。既是统计

认识为什么以原始信息为操作对象的理论前提,也是统计认识结果能够反映客体本质特征的理论依据。而作出这一假定的逻辑依据就是范式一。

范式二:

信息系统具有某种统计规律性

信息系统与客体系统在总体上同构

在总体上客体系统也具有该种统计规律性

这是将统计认识结果还原为客体本质特征的逻辑依据。

范式三:

某一样本具有特征(或遵从统计规律) $F$

该样本来自总体  $Q$

总体  $Q$  具有特征  $F$

这是根据样本推断总体的逻辑。

范式四:

某一总体具有统计规律  $P$

个体  $A$  来自该总体

$A$  也具有特征  $P$

这是解释总体的逻辑。

范式五:

在观测值范围内,被解释变量与解释变量之间具有某种统计模型

用于预测的解释变量在观测值范围内

被解释变量也服从某种统计模型

这是模型预测的逻辑。

范式六:

$a$  是  $F$ ,

$F$  是  $G$  的概率为  $q$ ,

因此,以概率  $q$ ,  $a$  是  $G$ 。



范式七:

$a$  是  $F$ ,

$F$  是  $G$  的概率小于百分之三,

---

所以,几乎一定地(或很可能) $a$  不是  $G$ 。

范式八:

$a$  是  $F$ ,

$F$  是  $G$  的概率极小,

---

所以,几乎一定地(或很可能) $a$  不是  $G$ 。

范式六、七、八仅在参考类  $F$  为有效时才可应用。因为只有这样,“ $F$  是  $G$  的比例”这句话才有明确的意义。但是,类似的不受这一限制束缚的论证类型则可由统计概率的频率解释来提出。在数学理论中,统计概率是由某些公理支配的集合测度;而一个描述集合  $G$  关于集合  $F$  的统计概率的形为  $P(G, F) = r$  的公式,粗略地说,就是  $G$  和  $F$  的交的测度除以  $F$  的测度等于  $r$ 。在应用于经验事物时,则是通过统计概率的频率解释而实现的。后者认为  $P(G, F) = r$  陈述了就全过程而言的相对频率  $r$ ;至于由人或自然界所进行的属于某个指定种类  $F$  的一次“随机试验”,则倾向于以该相对频率  $r$  产生种类  $G$  的结果。对于  $r$  接近于 1 的情形,这一频率解释通常表述为如下形式:如果  $P(G, F)$  非常接近于 1,那么只要种类  $F$  的一次试验刚一完成,实际上就一定会出现种类  $G$  的结果。实际上相当于如下推理格式,其中第二前提不再要求参考类  $F$  是有限的:

范式九:

$a$  是  $F$ ,

$F$  是  $G$  的统计概率接近与 1

---

所以, $a$  是  $G$  是几乎一定的

范式十:

$a$  是  $F$ ,

$F$  是  $G$  的统计概率接近于 0

所以,  $a$  是  $G$  几乎是完全不可能的。

#### 4.2.2 统计推理的不彻底性

事实上,所有的统计三段论都会导致不彻底性。因为结论指定给  $G$  (或,结论赋予它  $G$  的特征或属性)的个别事例  $a$ ,事实上可属于不同的参考类  $F_1, F_2, F_3, \dots$ ,而这些参考类的元素各以不同的相对频率或统计概率呈现为  $G$ 。例如,对于范式六式来说,假设张三是深圳人,而百分之九十九的深圳人是富人:但是张三又是一个教师,而百分之九十九的教师是穷人。于是根据推理格式范式六可以顺理成章地得出两个统计三段论:

张三是深圳人

深圳人是富人的比例为百分之九十九

因此,以概率百分之九十九,张三是富人

和

张三是教师

教师是穷人的比例为百分之九十九

因此,以概率百分之九十九,张三是穷人

上述两个推理的前提都是真的,但是,两个推论却根本不相容。

对于范式七来说,假设李四是大学一年级学生,大学一年级学生中年龄小于 17 岁的比例小于 2%,所以,李四的年龄几乎一定不小于 17 岁。而李四曾经跳过级,跳过级的学生有 2% 的年龄不大于 17 岁。所以,李四的年龄几乎一定小于 17 岁。同样在前提为真的条件下得出两个不相容的结论。相似地可以举出范式八的例子。

对于范式九,假设一只坛子里装了 1000 个球,其中,900 个是玻璃做的白球,1 个是象牙做的白球,99 个是象牙做的黑球。令  $E$  为从这只坛子中按照随机原则抓出一个球的程序。则作为

一次试验  $E$  的结果为白球的统计概率为  $p(\text{白}, E) = 0.901$ . 于是, 对于某一次具体抽球( $b$ )来说, 根据范式九, 得出如下推论:

$b$  是  $E$

$$P(\text{白}, E) = 0.901$$

因此,  $b$  是“白”(即该次抽出一个白球)是几乎一定的。

现在, 假定在  $b$  次抽取中, 得到一个水晶球, 则, 这时,  $b$  是红球几乎是一定的。

可见, 统计三段论从形式上看似乎有理, 其实却是站不住脚的。原因在于它会造成归纳的不一贯性。即对于一个具有统计三段论形式而前提为真的论证, 一般来说存在着一个具有相同形式而前提也为真的对立论证, 其结论与前一论证在逻辑上是不相容的。这一点对于另一种稍于区别的归纳规则也是对的。这一规则 *Black* 表述为:

$R$ : 根据在各种各样的条件下所考察的  $A$  的绝大多数事例是  $B$ , 从而推论(很可能)下一个遇到的  $A$  将是  $B$ 。

这里,  $R$  所支配的归纳论证, 其“强度”按照前提中所提供的有效事例的数量和多样性不同而不同; 因此, “虽然  $R$  允许我们断定某个无条件的结论, ……但断定的强度却随着证据的性质而波动”(洪谦:《逻辑经验主义》上卷, 商务印书馆, 1980 年版)。可见, 尽管规则  $R$  导致的结论不包含象“很可能”、“一定地”之类的情态修饰词, 但该结论却被设想为以或大或小的“强度”来断定的。因此, 按照规则  $R$ , 一个论证从真的前提导出一个对给定结论很强的断定, 一般来说会伴随一个遵从同一规则的对立论证, 该论证从同样为真的前提导出一个对于相反结论的强断定。同样具有不彻底性。

当然, 纯粹意义上的演绎推理形式决不会造成不一贯性, 特别是对于三段论形式的论证:

$a$  是  $F$

所有  $F$  是  $G$

$a$  是  $G$

这一论证的前提如果为真,就不会存在相同形式的对立论证——其前提也为真而结果却与已给论证的结论在逻辑上不兼容:不相容的结论只能从不相容的前提集合中演绎出来,而真前提的集合是不会不相容的。

## 4.3 统计推理

### 4.3.1 统计推理一般

统计推理,就是由一个或几个已知的统计判断,推导出另一些统计判断的思维形式。显然,其特点在于推理的构成要素都具有统计意义:作为推理依据的前提是统计判断,也就是说,前提本身就含有一定的不确定性;推理规则是统计法则;推理手段是统计方法,要得到任何一个统计结论,都必须进行大量的统计计算;推理的结论也是一个统计结论,具有一定的不确定性。

统计推理根据推理所依据的前提与结论的范围不同,可以具体化为五种推理形式。

1. 统计直接推理。是从总体到样本的推理,也称之为内在的推理或向下的推理。即:总体 $\rightarrow$ 样本(或个体)。例如,已知总体的分布推断样本和个体;根据已知的总体模型预测在解释变量的特定条件下被解释变量的对应值等都是直接推理。但是,值得注意的是,由于总体分布的不确定性,因而这种具有演绎形式的统计直接推理也是一种或然推理。

2. 统计逆推理。是从样本到总体的推理,也称向上推理。即:样本 $\rightarrow$ 总体或全域。例如:用样本模型代替总体模型、根据样本信息对总体参数或其他性质进行真伪的检验判断——假设检验以及利用样本信息对  $Y$  进行数量上的描述判断——参

数估计等都是逆推理。以极大似然法为例,它将所估计的总体参数  $Q$  视为变量,将样本统计量  $X$  视为已知量,寻找  $Q$  中具有最大可能产生样本观测值  $X$  的  $\hat{Q}$  作为  $Q$  的估计值。这种推理思想的根据是:由于样本是来自总体的,因而样本应该很好地反映总体的特征。极大似然估计正是根据样本特征寻找最可能产生样本的总体参数。

3. 统计比较推理。是根据它们之间已知的相似性从一个个体到另一个个体的推理。

4. 统计预报性推理。是从一个样本到另一个不同的样本的推理,也称外在推理。具有这种推理形式的最典型的统计行为就是统计预测。事实上,它是直接推理与逆推理的结合。即:样本 $\rightarrow$ 总体 $\rightarrow$ 样本(或个体)。其中,用样本回归模型(公式)代替(估计)总体理论模型(公式):样本 $\rightarrow$ 总体,是逆推理;而用模型值作为实际的预报值:总体 $\rightarrow$ 样本,是直接推理。

5. 全域推理。是从某个样本到全域形式的某一假设的推理。

统计的整个分析过程,同样是演绎推理与归纳推理相互依存的过程,但由于统计认识对象往往是随机的,或者是受许多偶然因素影响的复杂系统,也可能是原始信息是随机采集来的,因而不确定性是统计推理过程的一个显著特点。统计推断总是由不确定到比较确定或相当确定逼近的。

事实上,统计推理和经典演绎推理是一个整体,具有绝对确定性的演绎是作为概率的极限情况而得出的,于是,推理可表示为:

$$\begin{array}{ll} p=0 & h \text{ 与 } a \text{ 有矛盾} \\ a/h=p & 0 < p < 1 \quad \text{从 } h \text{ 依信度 } p \text{ 推出 } a \\ p=1 & \text{从 } h \text{ 能必然地推出 } a \end{array}$$

即在前提  $h$  下以合理信念度  $p$  相信结论  $a$ 。强调概率所表征的是推理过程,是两个命题(集)  $h$  与  $a$  之间的关系。

### 4.3.2 统计演绎

演绎推理,也称必然性推理。结论必蕴含在前提之中,这是一种封闭式思维。演绎推理是解释和预见原始信息、提出假说的重要方法。人们在发现了某些原始信息之后,需要运用演绎推理对其作出合理的解释,为科学预言和假说的提出指明正确的途径。科学史上的重大发现发明,有许多是先在理论上提出预言和假说,然后再在实践中得到证实的。在检验假说中,演绎推理起着十分重的作用。特别是许多假说难于直接在观测实验中得到检验,常常需要运用演绎法从假说中作出推论,预言还没有被人发现的事物和现象,再用观测实验来检验这些预言,从而使假说得到检验和发展。

统计检验就是一种演绎推理。统计检验的逻辑结构可以抽象为:

一般说来小概率事件在一次试验中不宜发生

某一事件是小概率事件

---

一次试验中该事件发生的可能性很小

然而,统计检验的逻辑不应认为是一种证明的逻辑,而是一种检验的逻辑。由于在统计检验的逻辑中,前提与结论之间的逻辑蕴涵不再是必然的,而是一种概率蕴涵,因而作为演绎的逻辑证明功能大大弱化,转化成为一种“证据支持”功能。也就是说,统计检验只能依据现有信息对原假设提供一定程度的支持,而不能逻辑地判“真”或断“伪”。在统计检验中所作出的“接受”和“拒绝”的判断,依据的是或然逻辑,而不是必然逻辑。“接受”中的“纳伪”和“拒绝”中的“拒真”的事实充分说明了这一点。

统计演绎具有一定的局限性。首先,统计演绎是一种创造性比较小的思维方法。其重要作用在于逻辑检验、或逻辑证明,而不在于科学发现。因为统计演绎是从一般到个别的推理,不可能对科学知识作出新的概括,所提供的新知识极为有限,不可

能用它总结出更普遍的科学原理,其结论原则上不可能超出前提的范围。其次,统计演绎推理的结论的可靠性受到前提的制约。由于演绎推理结论的正确性取决于前提的正确性,而前提是否正确,在演绎范围内是无法解决的。这就必须依赖归纳法和其它科学方法得出一般原理作为演绎前提。而归纳结论又有其或然性。更为突出的是统计演绎的前提除了科学规律以外,往往是统计规律,所以演绎结论也并不是绝对可靠的,也要靠归纳事实加以检验。因此,统计归纳和统计演绎客观上存在着内在的联系。演绎必须以归纳为前提,归纳也要以演绎为指导,二者相互依赖、相互渗透和相互促进。最直观的例子就是统计估计与统计检验之间的关系,二者密不可分,在相互作用中使统计认识不断接近客观真值。

#### 4.3.3 统计归纳

正如培根指出的那样,亚里士多德的三段论不是唯一的认识方式和推理方法。人们认识真理有两条途径:一是从一般到特殊(演绎);一是从特殊到一般(归纳)。而且后者比前者更容易获取真理。归纳是一种从个别概括出一般的思维方法和推理形式,是一种分析基础上的综合;它综合的是类对象某一方面或某几方面的共性,是一种高级求同分类的过程,演绎法从归纳法的结论开始,从一般推出个别。这是类中个体分解的分析过程。显然,归纳和演绎是分析综合的把个别上升为一般,把一般引申为个别的方法的一个特殊形式。归纳推理,也称或然推理。结论没有完全蕴含在前提之中,因此,即使前提为真,推理形式正确,结论只能或然地为真,有或大或小的概率。而概率的大小,又依赖前提给予结论支持的强度或确证程度。归纳概率就是用来表示归纳前提给予归纳结论支持的强度或确证程度的概念。

由于或然推理的结论不是直接地由前提导出,而往往是深入到前提所表征的事物内部或超出前提的集合范围,因而这种

推理常常伴随着对规律的认识,常常带来科学的新发现。在开放系统中向更深入更广阔探索。归纳推理又分为简单枚举归纳推理和科学归纳推理。简单枚举是根据某类事物的部分对象具有(或不具有)某种属性,而没有遇到相反事例,从而推导出该类事物都具有(或不具有)某种属性的推理。这种推理的依据是:某种共同属性在同类对象中不断重复,又没有遇到相反情况。这一推理对于获得一般性结论是必要的,但却不是充分的。因为结论所推断的超出了前提断定的范围。科学归纳推理是通过对某类事物的部分对象进行科学分析,找出它们与某种属性之间的因果关系或其他必然联系,从而对该类事物作出一般性结论。依据的是:事物与属性具有某种内在联系。总之,归纳是从个别到一般的推理方法。客观事物的个性中都包含着共性,通过个性可以认识共性,从个别到一般。科学认识的发展过程,需要从积累大量的观测、实验资料到概括一般原理。归纳法还是用来整理原始信息、从中得出普遍的规律或结论的一种基本方法,从已知推知未知。它以已知的关于原始信息的信息作前提,概括、解释新的原始信息,扩展认识成果,形成新的一般原理。

《概率论》和《统计学》等学科的发展,对传统归纳方法的变革产生了重大影响。把这些学科的成果引进归纳逻辑中来,用数理逻辑的工具对归纳推理予以系统的形式化和定量化,导致归纳概率和统计推理的形成和发展,丰富和拓展了归纳推理的内涵和外延。传统归纳理论主要研究确定的事件和现象,在很大程度上必须以机械决定论为其理论依据。然而,在多重因果联系中,特别是对于大量的随机现象来说,其运动规律则遵循概率规律,包括大数现象的规律性和具有随机特征的单个客体行为的规律性两个方面。这两类概率规律的概括同样要用归纳方法,但不是用传统的归纳法,而只能用统计方法。因此,可以说,概率规律是一种统计的规律,而统计方法则是揭示随机现象规律性的一种重要方法。



如果从一组归纳前提中,推导出几种结论,用归纳概率推算出其中每个结论的概率大小,就可以分别对它们的可靠性程度作出判断。这无疑有利于提高归纳概率的可靠性程度,但定量地解决归纳结论的确证度问题,虽有益于克服把归纳推理简单化等倾向,但其工作却是相当复杂、困难的。

统计推断首先要将样本资料提炼成有用的信息,这些信息主要以样本统计量的形式表现出来,具体表现为样本的数量特征值(均值、方差等)、样本数量关系(相关系数)及数量变化(回归方程),并由此对总体数量规律性做出统计推断。即

样本资料→样本信息→总体信息

统计方法不是纯粹的归纳:由结果到原因、由观测到结论、由样本到总体、或由局部到一般的统计结论,而是借助于在数学(概率论)中深入研究过的严格规则实现的。

统计方法还是在认识的经验层次上的即在对经验材料的观测、实验和直接加工的层次上的归纳和演绎的某种统一。由于采用了概率论,统计推理成为归纳法的精确化了的形式。

统计推断作为归纳推理的一种特殊形式,是对传统归纳法的丰富和发展。运用统计推理必须有正确的推理前提(依据),其前提应是对大量随机现象进行多次实验或观测得到的数据。其推理方式主要是运用数理逻辑,对作为推理前提的观测、实验数据进行估计、判断和推理,研究它们的频率、分布、各种平均量和各种类型的偏差等,然后概括出统计规律的一般结论。但是统计推断又不同于一般归纳,它以概率作为判据,而不要求若干个个别均具有某种特征,只要求以较大的概率具有某一特征即可。统计推断不仅扩展了归纳推理的研究范围,而且可以进行定量描述,成为归纳推理的一种精确化的重要形式。

统计归纳推理和传统的(普遍的)归纳法一样都是从部分(样本)的知识向整体(类、群、集)的知识的过渡。但是,统计推断与归纳概括的区别就在于:1. 统计推断确实也从若干个例(样

本)出发去追求对于总体(通常基数很大,甚至无限)的刻画,但它并不追求总体内每个成员共同具有的某种性质,并不企图建立全称命题形式的归纳概括。统计推断把总体当作一个独立的完整的对象来处理,描述和刻画总体本身的性质。这种性质对于总体单位来说常常是没有意义的。2. 从样本到得出关于总体的结论,需要执行复杂的程序,应用非经专门训练不能掌握的一套技术,远不象归纳概括那样径情直遂。常常对总体某一种性质的认识仅仅一个结论是不够的,往往需要有一个对差异的刻画。3. 从特殊到一般的形态中,归纳问题的困难主要是“证明”;而在统计形态中,归纳问题的困难主要是“发现”,虽然结论得出后也有一个证明(或检验)的问题,统计不仅是“证明的逻辑”,更是“发现的逻辑”。

此外,统计归纳推理还具体以下特点:1. 按照随机原则进行足够的枚举,尽可能使样本与总体同构,从而保证了样本(部分)对总体(全体)的代表性;2. 枚举中出现的反例提供了重要的信息,避免以偏盖全,使得对统计认识结果的理解和运用更科学、更客观,克服极端、盲目的倾向;3. 统计推理是一种数值推理,推理手段是数理逻辑,每一步推理都要进行大量的统计运算;4. 统计归纳是一种多值逻辑。传统的归纳是一种非此即彼的二值逻辑,即“非真即假”或“非假即真”。统计归纳与之不同,它用0到1之间的连续标度取代“真”和“假”这两个逻辑值。因此,人们认识的目的是追求“真”或“假”,而是接近客观存在的概率。在整个统计推理中,最低层次对于个别命题,仍保持二值性,往上一层对于命题序列,由其个别元素的真假值导出概率度。统计认识结论是一个或然判断。

#### 4.3.4 统计类比

统计领先指标法、统计模型法以及试验设计等都是类比推理在统计中的具体应用。所谓类比,就是根据两个(或两类)客

体系系统之间在某些方面的相似或相同,从一个客体或客体系统的知识向另一个客体或客体系统的知识过渡,推出它们在其他方面也可能相似或相同的一种逻辑推理方法。从内容上说,它既包含从特殊到特殊,又包含从一般到一般的推理方法。就推理过程而言,它从个别开始,以一般为中介,然后再达到个别。它是归纳和演绎的综合使用形式。

根据类比对象已知属性之间关系的不同。可把类比区分为各种不同的类型。其中主要有以下五种类型。

(1)简单共存类比。即只根据对象的属性之间有简单共存关系而进行的推理,并不知道属性之间是否还有其他关系。这种类比在很大程度上是一种预感或猜测,其可靠程度较低。

(2)因果类比。即根据两个对象的各自属性之间都可能具有同一种因果关系而进行的推理。在这种类比中,对象的各属性不再是简单共存的关系,而是有了某种因果的必然联系,因而提高了结论的可靠性。但这种因果关系终究是特殊对象的东西;不一定恰好适合另一特殊对象,因而它仍然是一种或然性的推理。

(3)对称类比。即根据对象的属性之间具有对称性而进行的推理。对称类比的结论往往比因果类比的可靠性程度要高一些。因为在一事物中与一方对称的另一方只有一个。但一事物中的对称关系,也不一定恰好适合另一特殊对象,所以它也还是一种或然性推理。

(4)协变类比。即根据对象的属性之间,具有某种确定的协变关系,即函数变化关系而进行的推理。协变类比定量地描述了对象的属性之间的关系,比前面几种只作定性描述的类比前进了一步,使对象的属性之间的关系更为确定,可靠性也更高。事实上,统计模型方法就是一种协变类比。它是根据数学形式的特征推断事物的特征。

(5)综合类比。即根据对象属性的多种关系的综合相似而

进行的推理。由于综合类比以属性的多种相似关系为依据,所以结论基本可靠。但实际上不可能把对象的所有关系都综合进去,因此仍然具有或然性。

类比法在各种逻辑推理方法中,是最富于创造性的一种方法。首先,其应用范围广泛。如前所述的归纳法和演绎法,其应用范围只局限于同类事物的联系。而进行类比的两个事物可以是同类的,也可以是不同类的,有利于人们充分发挥想象力,在广阔的范围内把不同事物联系起来进行类比。促进具有创造性的统计认识活动。其次,类比法的应用不受已知知识的束缚。人们进行统计认识活动,在借助于已知知识由已知推测未知过程中,演绎法通常会受到作为前提的一般原理的限制,归纳法则过分地受到特殊知识数量上的限制。因此,当发现了某些原始信息之后,由于旧的理论无法解释,演绎法对此无能为力,由于原始信息数量太少,归纳法也无从下手,这时,类比法可以只根据少数特殊知识,进行具有创造性的认识研究,提出科学的假设和推断,起到探索、预测的作用。就象康德所说的那样:“每当理智缺乏可靠论证的思路时,类比这个方法往往能指引我们前进。”(康德,《宇宙发展史概论》,上海人民出版社,1972年版,第147页)

但是,必须指出的是,类比法在各种逻辑推理方法中,其可靠性最小。也就是说,运用类比法,常常会推导出错误的结论。之所以如此,主要有以下三方面的原因:第一,是由类比法的客观基础所决定的。类比法实质上是异中求同的方法。客观事物之间既有同一性,又有差异性,客体之间的相似性或同一性,使类比法有可能获得正确的结论;而客体之间的差异性,又使类比法的结论带有不正确性。如果根据两个事物具有的相似性进行类比推理,推出的属性正好是它们的差异性时,类比法的结论就会发生错误。第二,类比推理的逻辑根据是不充分的。类比法是以两个对象的某些属性相似或相同为前提,推出他们在其它

属性方面也相似或相同的结论。其前提和结论之间并没有必然的联系,只是一种可能性。这就决定了类比只是一种或然性推理。第三,类比法的推理规则是很不严密的。逻辑推理的结论的可靠性似乎与其推理规则的严密程度成正比。演绎法具有最严密的推理规则,归纳方法次之,类比法最小。从认识功能来看,类比法的可靠性虽最小,而其创造性却最大。逻辑推理的可靠性虽与其推理规则的严密程度成正比,却又似乎与其创造性成反比。演绎法虽具有最严密的推理规则、可靠性最大,但创造性却最小。而类比法则相反。由于它的逻辑根据不充分和推理规则不严密等原因,使得其可靠性最小,但其应用范围宽广,又不受已知知识的限制,而且又具有最大的创造性。因此,在应用类比法时应扬长避短。

#### 4.3.5 贝叶斯推理

##### 一、贝叶斯统计的基本思想

贝叶斯定理的原始形式为:在成功  $P$  次,失败  $q$  次的情况下,  $X \in (a, b)$  的概率为:

$$\frac{\int_a^b X^P (1-X)^q dx}{\int_0^1 X^P (1-X)^q dx}$$

贝叶斯曾对上述定理给出了如下的几何解释。下图是边长为 1 的正方形桌子  $ABCD$ ,设想一个被蒙住眼睛的人往桌子上随意掷球,球落在桌子上各部位的机会均等,这一点蒙眼人知道,但他当然不知道球落点的确切位置  $W$ 。过  $W$  作  $EF$  平行于  $AB$ ,从而把正方形分成两部分。然后让蒙眼人多次掷另一个球,如落在  $EF$  的右边算成功,落在左边算失败,告诉他成功了  $P$  次,失败了  $q$  次,蒙眼人的任务是对  $EF$  的位置(即  $AF$  的长度  $X$ )作出估算。按贝叶斯的算法,就是求  $X \in (a, b)$  的概率为多少。容易定性地推断,成功的次数越大,失败的次数越小,  $EF$  的

位置越靠左;反之,越靠右。极端情况下,全数成功,聪明的蒙眼人会推测出  $EF$  几乎与  $CD$  重合;全数失败,则  $EF$  几乎与  $AB$  重合。

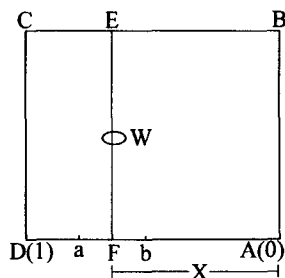


图 14 贝叶斯定理几何示图

常用的 Bayes 定理:

$$P(B/A) = P(AB)/P(A)$$

后人常把未作任何试验时事件  $M$  的概率叫做先验概率,相应地把已有经验证据条件下推算的概率叫做后验概率。其实,在 Bayes 的思想中,并没有这种能与“先验概率”相提并论的“后验  $P$ ”。只有“先验  $P$ ”是存在的、确定的,但只是对无限的种而言;有限的人充其量只能根据经验作出某种推测,因此,后验概率并非对先验概率的修正,而是人在无可解脱的束缚下追求先验  $P$  唯一可行的方式,是对先验概率的后验测算。这种追求永远达不到目的。因此,人只能限于了解先验概率位于某一区间的可能性有多大。

已经观测到 100 只白天鹅,就算进行了 100 次全数成功的试验,照 Bayes 定理计算出的并不是第 101 只天鹅为白的概率,而是“天鹅皆白”这一无限全称陈述的概率。(关于下一次试验的成功概率,拉普拉斯给出了另外的公式)。普赖斯指出:Bayes 定理“提供给我们必要的指针,以便从效果确定未知原因的本性与其作用的大小,他又提供给我们强有力的防护以摆脱哲学家

常常陷入的一个巨大危险。这个危险来自于,一方面把结论建基在不充分的归纳上,另一方面,又要求结论能得到比单纯试验次数所提供的更强的担保”。(The History of Statistics of 17 and 18 Centuries 第 370 页)

## 二、贝叶斯统计的认识逻辑

根据贝叶斯的概率公式:

$$P(A_i/X) = \frac{P(A_i)P(X/A_i)}{\sum P(A_i)p(X/A_i)}$$

我们不难看出如下认识。

1. 贝叶斯公式既考虑主观概率又尊重客观信息。在进行统计认识之前,关于客体的知识可能来自某种理论、以往经验、主体的分析判断等,这是宝贵的财富,是人类对客体认识的结晶,毫无疑问应当充分利用。但是,毕竟是过去,毕竟是经验,是信念。公式中分子的第一项  $P(A_i)$  是未知参数  $A$  的先验概率(验前概率),一般在本次试验前,根据以往的经验(认识)加以确定的。第二项  $P(X/A_i)$  是以参数为条件,通过样本资料得到的新信息。 $P(A_i/X)$  是计算出来的后验概率,它是以样本  $X$  为条件,是先验概率和样本信息的综合。把经验与现实有机地统一起来。同时,在公式中很好地体现了统计的历史观:任何系统都具有“记忆性”;系统对这种“记忆性”的记忆程度将随着系统的稳定性而有所不同;对于渐近稳定系统来说,系统的“记忆性”将随着时间的推移而淡忘(递减);对于边界稳定系统来说,系统的“记忆性”是个恒定量即既不“忘却”,也不“增强”;对于非稳定系统来说,系统的“记忆性”则无限增大。

2. 强调认识是一个过程,是认识的积累。本次得到的后验概率是在以往所有的认识基础上得到的,是动态概率,随着信息(试验次数)的不断增加,统计认识不断深化。

3. 根据对客体的把握程度,具体设计试验,后验分布的均值是先验分布的均值( $X$ )和样本均值( $\bar{X}$ )的加权平均,即  $Y =$

$W_1 X + W_2 \bar{X}$ 。权数  $W_1$  和  $W_2$  的确定并非唯一,视先验和样本信息的可信程度而定。如果认识主体对先验概率有较大把握(比如说,经过了多次试验修正,有近期试验,有一定的理论支持),则可将  $W_1$  取大一点,  $W_2$  小一点;否则,  $W_1$  取小一点,  $W_2$  取大一点,以加重样本信息的作用和地位。就样本的设计来说,若对先验概率有较大的把握,则样本可小一些,主要是利用样本信息对我们验前经验做一个验证和补充;若对先验概率把握不大,则应取较大的样本,以平抑先验概率的影响,而增加样本信息的作用。

4. 贝叶斯统计概率是一个由过去到现在,由主观到客观的一个认识过程,随着样本信息的增加,初始确定的主观概率的不合理性逐渐减弱,而由客观概率代之。

5. 贝叶斯统计是实践——认识——再实践——再认识的循环过程。从贝叶斯公式直观来看,推断是从先前的认识开始的,在现有的认识的基础上,进行试验,即实践,然后得出对客体更深刻的认识,即遵循着认识——实践——再认识的过程。但是,这是就某一次认识而言的。追溯其初始的先验(主观)认识,不论是哪一种获得先验分布(初始概率)的方法,都是源于实践的,比如,先通过某种调查方式,获得一个分布比律;共轭分布族法则是从样本分布出发,由样本概率密度推导出先验分布。广义贝叶斯估计是根据样本信息将极大似然估计、最小二乘估计等作为贝叶斯的极限估计来代替;经验分布法的“经验”来自于对实践观测的历史资料,就是同等无知原则,看起来似乎是主观的,实际上它相当于把该次试验的样本信息作为首次初始(先验)信息。因为对先验概率给予均匀分布,就相当于没有先验概率。此时,样本信息直接影响后验概率。

### 三、贝叶斯统计认识的实例分析——《黔之驴之统计诠释》

《黔之驴》这是一则众所周知的寓言故事,写的是一头驴被老虎吃掉的故事。从统计认识的角度来看,是老虎不断获得信



息,深化认识,修正行为,作出正确决策的过程,是贝叶斯“相继律”的应用过程。也就是说,毛驴这一“庞然大物”被老虎吃掉这一现象的背后,存在着统计认识的必然判断。下文试图对这一故事予以统计解读,以说明统计决策是一个不断获得信息,使认识无限逼近客观存在的认识过程,可以广泛地运用于各个领域。

“弱肉强食”这是动物界生存的一条基本法则。老虎也不例外。当它猎物时,对方有两种可能状态:或者是强于己——强者(记作  $Q$ ),或者是弱于己——弱者(记作  $R$ )。老虎的行动也有两种选择:一是攻击(记作  $G$ ),二是逃生(记作  $T$ )。因此,老虎要先对对方进行判断:与己相比,它是强者还是弱者。用统计认识的观点来看,就是老虎需要决策。在这一决策问题中,  $Q$  和  $R$  构成了该决策的状态空间,  $G$  和  $T$  是行动(策略)空间。而驴可能是强者,也可能是弱者的这种不定性,就是状态概率  $P(Q)$  和  $P(R)$ ,现在的问题就是尽可能获取信息(Data),作出正确的判断。

猎物者(老虎)的信息来自对方(驴)的行动。因为弱肉强食这一法则对任何动物都适用。假设就一般而言,确实是强者遇到较弱的猎物时,进攻的概率为 80%,不进攻的概率为 20%;而确实是弱者遇到较强的猎物时,由于错误判断或饥饿危及生命也会决一死战,因而也有可能发起进攻,假设其概率为 10%,不进攻的概率为 90%。于是,得到了 Data(驴的行为——进攻与否)的概率分布如下:

Data 的概率分布表

	进攻 $B_1$	没进攻 $B_2$
强者( $A_1$ )	0.8	0.2
弱者( $A_2$ )	0.1	0.9

老虎刚到贵州时,“虎见之,庞然大物也,以为神”。显然,这时老虎认为驴是强者的可能性较大,设其概率为  $P(Q) = 0.9$ , 认为驴是弱者的概率为  $P(R) = 0.1$ 。因此,未敢轻举妄动,而是“蔽林间窥之”。

为了获得信息,老虎“稍出近之”,“愁愁然,莫相知”,此举没有引起驴的反映,未获得任何消息,所以老虎对驴的认识仍未改变。

“他日,驴一鸣,虎大骇,远遁”。虽说此时是老虎逃走,驴暂时保全了性命,可驴并没有向老虎发动实质性的进攻。老虎得到了重要的信息,即事件  $B_2$  发生了,根据这一信息,老虎对驴的认识发生了变化:

驴是强者的概率为

$$\begin{aligned} P(Q_1) &= P(A_1/B_2) \\ &= \frac{P(A_1)P(B_2/A_1)}{P(A_1)P(B_2/A_1) + P(A_2)P(B_2/A_2)} \\ &= \frac{0.9 \times 0.2}{0.9 \times 0.2 + 0.1 \times 0.9} \\ &= 0.67 \end{aligned}$$

驴是弱者的概率为

$$\begin{aligned} P(R_1) &= P(A_2/B_2) \\ &= \frac{P(A_1)P(B_2/A_2)}{P(A_1)P(B_2/A_1) + P(A_2)P(B_2/A_2)} \\ &= \frac{0.1 \times 0.9}{0.9 \times 0.2 + 0.1 \times 0.9} \\ &= 0.33 \end{aligned}$$

这时,老虎仍然是“终不敢搏”。只好进一步获得信息。老虎“稍进,益狎,荡倚冲冒”。而“驴不胜怒”,但仅是“蹄之”罢了,仍未作出实质性的进攻。根据所得到的信息,老虎进一步修正了自己对驴的认识。

驴是强者的概率为:

$$\begin{aligned}
 P(Q_2) &= P(A_1/B_2) \\
 &= \frac{0.67 \times 0.2}{0.67 \times 0.2 + 0.33 \times 0.9} \\
 &= 0.30
 \end{aligned}$$

驴是弱者的概率为：

$$\begin{aligned}
 P(R_2) &= P(A_2/B_2) \\
 &= \frac{0.33 \times 0.9}{0.67 \times 0.2 + 0.33 \times 0.9} \\
 &= 0.70
 \end{aligned}$$

此时，老虎甚喜，对驴的本领有了较深刻的认识，得出“技止此耳”的结论。因此，采取了进攻的对策，即“跳跟大闹，断其喉”。由于老虎获得了足够的信息，经过不断地认识，使其认识逐渐接近客观，在此基础上，采取了正确的决策——“进攻”。因而，“尽其肉，乃去”，大获全胜。

这里有一个问题：为什么老虎在根据第一次获得的信息修正认识后，认为驴是强者的概率为 67% 时，没有进攻，而在根据第二次获得的信息修正认识后，认为驴是强者的概率为 30% 时，立即采取了“进攻”的对策呢？毫无疑问，是因为在第二次获得信息后，认为驴是强者的可能性较小了。那么究竟当概率为多大时，就可以做出判断，采取对策呢？也就是说，概率为何值时，就可以认为较大或较小呢？是否是以 50% 为界呢？

为此，我们来讨论决策的效率损失。设损失阵如下：

	进 攻	不 进 攻
驴是强者	$a$	$c$
驴是弱者	$b$	$d$

则，进攻的损失函数为期望损失：

$$L_1 = L(a, b) = P(A_1)a + (1 - P(A_1))b$$

不进攻的期望损失为：

$$L_2 = L(c, d) = P(A_1)c + (1 - P(A_1))d$$

显然,期望损失与状态概率和损益值有关:

当  $P(A_1)=0$  时,对于  $L_1$  来说,有  $L(a,b)=b$ ;对于  $L_2$  来说,有  $L(c,d)=d$ .

当  $P(A_1)=1$  时,对于  $L_1$  来说,有  $L(a,b)=a$ ;对于  $L_2$  来说,有  $L(c,d)=c$

于是,损失函数图如下:

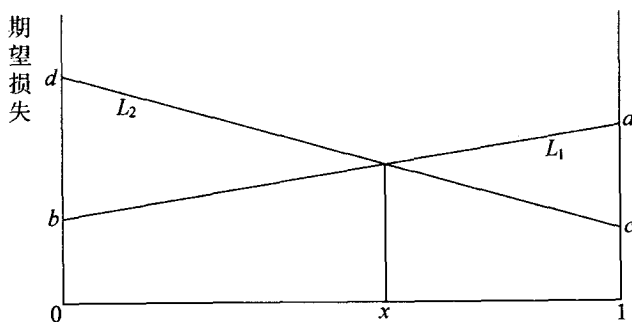


图 15 损失函数图

显然,  $L_1$  与  $L_2$  的交点  $x$  为决策的转移点,当认为驴是强者的概率在  $(0, x)$  范围内,即  $0 < P(A_1) < x$  时,进攻的损失小于不进攻的损失,显然应采取“进攻”的对策;当认为驴是强者的概率在  $(x, 1)$  范围内,即  $x < P(A_1) < 1$  时,不进攻的损失小于进攻的损失,显然,应采取“不进攻”的对策。

根据上述分析,交点  $x$  是根据认识采取对策的重要依据。不难看出,交点为:

$$x \cdot a + (1-x) \cdot b = x \cdot c + (1-x) \cdot d$$

解之,得

$$x = \frac{d-b}{(a-c)+(d-b)}$$

显然,交点  $x$  为损失值的函数。在理论上,它是在给定的损失函数下,后验分布的贝叶斯解。只有当  $a-c=d-b$  时,交点

$x$  才是 0.5(即 50%)。

假设,老虎采取两种对策的损失阵如下:

	进 攻	不 进 攻
驴是强者	100	20
驴是弱者	10	130

这时,交点

$$\begin{aligned}x &= (130 - 10) / [(100 - 20) + (130 - 10)] \\&= 120 / 200 \\&= 0.6\end{aligned}$$

由于老虎第一次修正认识后,认为驴是强者的概率为 0.67,因而,不宜发起进攻;而经第二次修正认识之后,认为驴是强者的概率为 0.23,小于交点 0.6,所以果断地采取了“进攻”的对策,并取得了胜利。

总之,从一般认识过程来看,《黔之驴》这则故事反映了老虎对一“不速之客”的认识过程。从统计认识来看,是一个典型的不断利用新的信息(Data)修正先验概率的统计决策认识过程。它告诉我们:统计认识是一个过程,不能寄希望于一蹴而就;先验认识的正确与否并不十分重要,至关重要的只是能否获得足够的客观信息,不论是由于没有任何先验信息,因而先验概率取等概分布也好,还是先验信息、先验概率失真也罢,只要不断获得真实客观的信息,就可逐渐修正先验概率,使之接近客观;在根据概率作决策时,要结合损益值来确定采取对策的转移点,要把握好作决策时所依据的统计认识程度,不应一味追求彻底认识即更大的概率值,可以设想,若老虎再进行几次试探,驴是强者的概率更小,这时再作决策,固然把握性更高,但获取更多的信息要付出成本,更为甚者,会因此而误失良机。统计认识方法适运用于人类的一切认识活动。

## 4.4 统计解释及其逻辑模型

### 4.4.1 统计解释

关于统计解释的思考前已述及,这里不再重述。

### 4.4.2 统计解释的逻辑模型

由于解释是将个别现象归入一般规律,因此,其基本逻辑要素是一般规律( $C$ )和个别现象( $E$ )。一般规律被限制在特殊条件下,从而表现为个别现象。

其基本逻辑模型为:

$$\begin{array}{c} L \\ C \\ \hline E \end{array}$$

可简化为  $L \wedge C + E$

在这里,  $L$  起解释大前提的作用。只有规律所涵盖的个别才是可解释的,也只有规律才授予先行条件以解释者的资格。但是单从一般规律是不能演绎出关于被解释现象的陈述的。因此,解释前提除一般规律外,还必须包括导出被解释现象所必需的那些先行条件。通常包括:使一般规律发生作用的边界条件;发生在被解释事件之前或同时发生的初始条件。一般说来,解释一个简单现象只需一个一般规律和先行条件就够了。但是要解释比较复杂的现象,就需要多个规律和多种条件,表现为科学规律集和先行条件集。因此,解释的一般模型为:

$L1, L2, \dots, Ln$  (规律语句集)

解释者

$C1, C2, \dots, Cm$  (先行条件语句集)

被解释者       $E$  (被解释者语句集)

或表示为： $\{Ln\} \wedge \{Cm\} + E$

这里，被解释语句  $E$  可以是单称的事实，即特定时空的特定条件；也可以是经验概括和经验定律；甚至还可以是理论定律。在这种一般模型中，如果未知项目（待发明项目）不同，那么解释图式也将有所差异，较常见的有以下几种情况：

(1) 对解释的大前提知之不多，需通过发现过程完善规律语句集，这种解释图式包括着新概念、新假说乃至新理论的发现，是一种重要的图式。

(2) 对规律起作用的条件知之不多，需通过发现过程完善先行条件集。这种解释包含着新事实、新关系、新结构的发现，可以把不同层次的规律联系起来。也是一种重要的解释图式。

(3) 对于如何从前提推演出结论的逻辑推导过程不清楚，需运用各种方法建立从前提到结论的逻辑桥梁，一旦找到了推进形式完善化之路，往往会成为新的理论发现。例如，运用统计平均方法处理分子体系，最初是为了将经典力学规律与热力学现象联系起来，尔后却发展为统计物理。

(4) 当解释前提已知，而被解释现象未知的情况下，解释模型变成了关于新事实的预言模型。有时关于新事实的预言与关于先行条件的知识，有一定程度的互换性，在这种情况下，新事实预言与先行条件发现之间并没有明确界限。

值得注意的是，在上述各种图式中，如果用以进行科学解释的定律不是全称普遍规律而是统计规律，那么，这种解释的一般形式为：

$$L: (x)(Px \rightarrow Pr(Qx) = r)$$

$$\frac{C: Pa}{E: Qa} [r]$$

可简化为： $L \wedge Cr + E$

这里的一般规律是：对于一个具有性质  $Px$  的任何客体  $x$ ，以  $Pr$  的概率出现性质  $Qx$ ，如果一个特殊客体  $a$  具有性质  $P$ ，则

以  $r$  的概率推出  $a$  呈现性质  $Q$ 。

这种解释所依据的规律不是全称规律,而是一种统计规律,具有以下特点:

(1)它只断言某种现象、某种联系以一定的或然性出现;

(2)解释者或解释陈述不是逻辑地蕴涵被解释陈述,因此在概率解释中,前提为真并不演绎地决定结论为真,它是允许反例存在的。

(3)概率解释在本质上是归纳的,这种归纳性质,来自客观事件的概率分布,当客观概率传递给解释过程时,解释本身也成为或然推理。



## 参考文献

- [1] 《马克思恩格斯选集》1-4卷,人民出版社 1972 年版。
- [2] 《马克思恩格斯全集》第 20 卷,人民出版社 1965 年版。
- [3] 《毛泽东选集》(一卷本),人民出版社,1964 年 4 月。
- [4] (英)罗素:《西方哲学史》,下卷,商务印书馆,1976 年版。
- [5] (荷兰)斯宾诺莎:《笛卡尔哲学原理》,商务印书馆,1979 年版
- [6] [苏]柯谱宁著,彭漪涟、王天厚等译,《作为认识论和逻辑的辩证法》,华东师范大学出版社,1984 年版
- [7] 《方法概念》《The Concept of Method》纽约和伦敦,1961 年
- [8] 陈选善:《教育测量》,商务印书馆
- [9] 黑格尔:《逻辑学》下册,商务印书馆
- [10] 费而斯曼:《门德列也夫周期律在现代科学中的应用》,中国青年出版社,1959 年版
- [11] 亚里士多德:《后分析篇》,见《工具论》,广东人民出版社,1984 年版
- [12] 康德:《宇宙发展史概论》,上海人民出版社,1972 年版
- [13] 《列宁全集》第 27 卷,人民出版社,1959 年版
- [14] 黑格尔:《小逻辑》,商务印书馆,1997 年版
- [15] [苏]拉扎列夫等:《认知结构与科学革命》湖南人民出版社 1986 年版
- [16] 维纳:《控制论》,科学出版社,1962 年版
- [17] 维纳:《人人有人的用处》,商务印书馆,1978 年版
- [18] 周概荣:《概率论与数理统计》,高等教育出版社,1984 年版

- [19] 茆诗松等:《高等数理统计》,高等教育出版社,1998 版
- [20] 邱东:《多指标综合评价方法的系统分析》,中国统计出版社,1991 年版
- [21] 佟哲晖、邱东:《国民经济统计学》,中国统计出版社,1992 年版
- [22] 王振龙:《时间序列分析》,中国统计出版社,2000 年版
- [23] 王振龙:《统计学原理》,陕西科学技术出版社,1993 年版
- [24] 岳巍:《统计认识论》,中国统计出版社,1994 年版
- [25] 袁卫:《统计推断思想》,中国统计出版社,1989 年版
- [26] 刘李胜:《决策认识论导论》,山西经济出版社,1995 年版

## 后 记

本书是根据我的博士论文《统计哲学思考》修改而成的。关于“统计哲学”这一概念,一般会有两种理解:一是统计的哲学属性;二是哲学在统计中的应用。前者研究的是统计作为人类认识工具的本质属性及其认识方法的普适性,属于哲学一般的范畴。而后者则是指统计作为哲学的一个认识领域,研究如何在统计中贯彻哲学思想和应用哲学方法的问题,在这种意义下,统计哲学类似于社会哲学、心理哲学、艺术哲学等属于应用哲学的范畴。我的研究聚焦于前者,即是对统计的哲学属性的探索,是对统计哲学学科体系构架的探索。

早在1991年我撰写的《对传统统计整理的反思》一文,从哲学的视角对统计整理方法和过程进行了剖析,随后相继发表了《论统计信息的特性》、《论统计信息的运动过程和统计信息系统》等文章,提出了统计活动的产品是统计信息的观点,讨论了统计信息概念的内涵和外延,论述了统计信息的特殊运动过程。1993年在《统计研究》期刊上向统计人发出了呼吁:《还统计学其本来面目——亚哲学》。1995年在国际统计学会第50届大会上作了题为《统计学的哲学思考》的演讲。1996年考取了东北财经大学的博士研究生,在邱东教授的指导下,攻读统计学,并选择《统计哲学思考》作为博士学位论文的题目。从此开始了在华山脚下登哲学书山,在渤海之滨游统计学海的探索生涯,经受了书山之坎坷,饱尝了学海的苦涩!今天终于得以小憩。长叹之余,艰难险阻尽入眼帘,咸酸苦辣涌上心头……。

没有半途而废就是胜利。因而,不禁一种征服者的感觉油然而生……

拙著今天得以面世,我首先要感谢的是我的导师邱东教授,在崎岖的书山中他给我指路,在茫茫的大海里他给我导航。我不仅从他那里学到了许多知识,更重要的是学到了治学的精神、探究的方法以及为人处世的原则。还要感谢蒋萍教授、白雪梅教授、王庆石教授、于洪平教授、王维国教授以及所有关心和帮助过我的恩师益友、前辈学弟。同时也不能忘记西安财经学院的各位同仁在学习上给予的启迪、工作上给予的支持、生活上给予的帮助,以及中国统计出版社的领导和员工为出版拙著而给予的支持和付出的劳动。

在这里更不能不提的是我的贤妻和爱子。她(他)们不仅包揽了全部家务,而且还要千方百计地调剂饮食的营养和口味、生活的节奏和情趣。贤妻王莉霞博士还对我的论文中的一些观点提出了许多宝贵意见;年仅13岁的爱子王超帮我完成了大量的手稿录入工作。

对于这一切的一切,我的感激之情难以言表,愿以拙著聊表回报!

这一课题研究要回答的问题很多,且很复杂,拙著只是一个阶段成果。从初衷来看,是想建立一门“统计哲学”,但是由于时间和精力所限,离较为系统、完善的一门学科尚有很大距离。不仅是我还需要做长期的、艰苦的、更深入的探索和研究,而且还需要更多的统计学人共同研究。从内容上来看,该课题是用哲学这面多棱镜全方位透视统计实践和统计科学,涉及的点多、面广,因而有许多问题只是提出观点和看法,而没有做详尽的阐释。从研究方法来看,是本着探索的精神对统计所作的一些思考,因此可能有些观点、提法、表述和阐释尚需作进一步的论证和讨论。所有这些都有待作更艰辛、更理性的探索,同时也期待着各位专家和读者指教。

王振龙 谨识

2001年秋于西安